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

Optimization problem in standard form

minimize
$$f_0(x)$$

subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $h_i(x) = 0, \quad i = 1, \dots, p$

- $x \in \mathbb{R}^n$ is the optimization variable
- $f_0: \mathbb{R}^n \to \mathbb{R}$ 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:

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

- $p^* = \infty$ if problem is infeasible (no x satisfies the constraints) $x^* = \infty$ if problem is infeasible (no x satisfies the constraints) $x^* = \infty$
- $p^* = -\infty$ if problem is unbounded below $\min x$ st. $x \le 5$

Optimal and locally optimal points

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

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

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

minimize (over
$$z$$
) $f_0(z)$ subject to
$$f_i(z) \leq 0, \quad i=1,\ldots,m, \quad h_i(z)=0, \quad i=1,\ldots,p$$
 $\|z-x\|_2 \leq R$

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

- $f_0(x) = 1/x$, $\operatorname{dom} f_0 = \mathbf{R}_{++}$: $p^* = 0$, no optimal point fo(x) -> 0 as x -> +inf
- $f_0(x)=-\log x$, $\operatorname{dom} f_0=\mathbf{R}_{++}$: $p^\star=-\infty$ f0(x) -> -inf as x-> +inf
- $f_0(x) = x \log x$, $\operatorname{dom} f_0 = \mathbf{R}_{++}$: $p^* = -1/e$, x = 1/e is optimal See f0(x) in [0,2]
- $f_0(x)=x^3-3x$, $p^\star=-\infty$, local optimum at x=1 See f0(x) in [-3,+3]

Implicit constraints

the standard form optimization problem has an implicit constraint

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

- ullet 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:

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

is an unconstrained problem with implicit constraints $a_i^T x < b_i$ iff b_i - a_i x > 0

Feasibility problem

find
$$x$$
 subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$ $h_i(x) = 0, \quad i = 1, \dots, p$

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

minimize
$$0$$
 subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$ $h_i(x) = 0, \quad i = 1, \dots, p$

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

Convex optimization problem

standard form convex optimization problem

minimize
$$f_0(x)$$

subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $a_i^T x = b_i, \quad i = 1, \dots, p$

- f_0 , f_1 , . . . , f_m are convex; equality constraints are affine
- ullet problem is *quasiconvex* if f_0 is quasiconvex (and f_1 , . . . , f_m convex)

often written as

minimize
$$f_0(x)$$

subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $Ax = b$

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

example

minimize
$$f_0(x) = x_1^2 + x_2^2$$

subject to $f_1(x) = x_1/(1+x_2^2) \le 0$
 $h_1(x) = (x_1 + x_2)^2 = 0$

- 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

to the convex problem

minimize
$$x_1^2 + x_2^2$$

subject to $x_1 \le 0$
 $x_1 + x_2 = 0$

Local and global optima

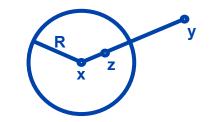
any locally optimal point of a convex problem is (globally) optimal

proof: suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$ (i.e., x not globally optimal) not in the local region

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

$$z$$
 feasible, $||z - x||_2 \le R \implies f_0(z) \ge f_0(x)$

then θ |v-x| 2 = R/2



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

not in the local region

$$||y-x||_2 > R$$
, so $0 < \theta < 1/2$

ullet z is a convex combination of two feasible points, hence also feasible

which contradicts our assumption that x is locally optimal since we found that fo(z) < fo(x)

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

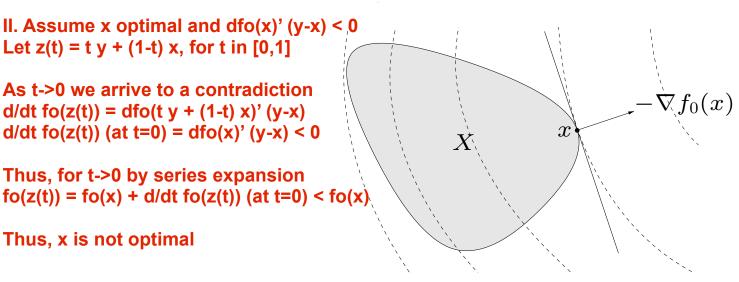
1st order condition for convexity
$$fo(y) \ge fo(x) + dfo(x)'(y-x)$$

$$\nabla f_0(x)^T(y-x) \ge 0$$
 for all feasible y

I. Assume $dfo(x)'(y-x) \ge 0$ then $fo(y) \ge fo(x)$ then x optimal

Let z(t) = t y + (1-t) x, for t in [0,1] As t->0 we arrive to a contradiction d/dt fo(z(t)) = dfo(t y + (1-t) x)' (y-x) $d/dt \ fo(z(t)) \ (at \ t=0) = dfo(x)' \ (y-x) < 0$ Thus, for t->0 by series expansion fo(z(t)) = fo(x) + d/dt fo(z(t)) (at t=0) < fo(x)

Thus, x is not optimal



(If it were an unconstrained problem the optimal x would be in this region.)

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

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

minimize
$$f_0(x)$$

subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $Ax = b$

is equivalent to

minimize (over
$$z$$
) $f_0(Fz+x_0)$
subject to $f_i(Fz+x_0) \leq 0, \quad i=1,\ldots,m$

where F and x_0 are such that

$$Ax = b \iff x = Fz + x_0 \text{ for some } z$$

• introducing equality constraints

minimize
$$f_0(A_0x + b_0)$$

subject to $f_i(A_ix + b_i) \leq 0, \quad i = 1, \dots, m$

is equivalent to

minimize (over
$$x$$
, y_i) $f_0(y_0)$ subject to $f_i(y_i) \leq 0, \quad i=1,\ldots,m$ $y_i=A_ix+b_i, \quad i=0,1,\ldots,m$

introducing slack variables for linear inequalities

minimize
$$f_0(x)$$

subject to $a_i^T x \leq b_i, \quad i = 1, \dots, m$

is equivalent to

minimize (over
$$x$$
, s) $f_0(x)$ subject to $a_i^T x + s_i = b_i, \quad i = 1, \dots, m$ $s_i \geq 0, \quad i = 1, \dots m$

• epigraph form: standard form convex problem is equivalent to

minimize (over
$$x$$
, t) t subject to
$$f_0(x) - t \leq 0$$

$$f_i(x) \leq 0, \quad i = 1, \dots, m$$

$$Ax = b$$

• minimizing over some variables

minimize
$$f_0(x_1,x_2)$$
 subject to $f_i(x_1) \leq 0, \quad i=1,\ldots,m$

is equivalent to

minimize
$$\tilde{f}_0(x_1)$$
 subject to $f_i(x_1) \leq 0, \quad i = 1, \dots, m$

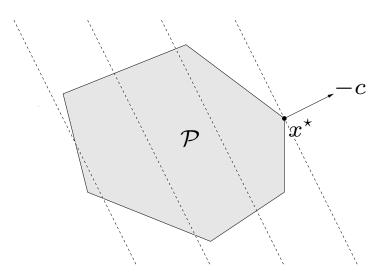
where
$$\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$$

Linear program (LP)

minimize
$$c^T x + d$$

subject to $Gx \leq h$
 $Ax = b$

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



Examples

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

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

to find cheapest healthy diet,

minimize
$$c^T x$$
 subject to $Ax \succeq b$, $x \succeq 0$

piecewise-linear minimization

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

equivalent to an LP

minimize
$$t$$
 subject to $a_i^T x + b_i \leq t, \quad i = 1, \dots, m$

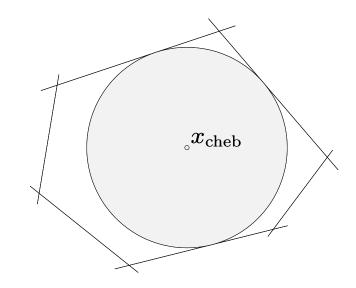
Chebyshev center of a polyhedron

Chebyshev center of

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

is center of largest inscribed ball

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



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

$$\sup_{\mathbf{u}} \{a_i^T(x_c + u) \mid \|u\|_2 \le r\} = a_i^T x_c + r \|a_i\|_2 \le b_i$$
 since sup **u** (a i u) = |a i| 2 by norm duality

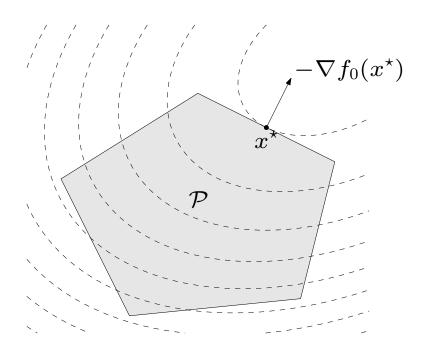
ullet hence, x_c , r can be determined by solving the LP

maximize
$$r$$
 subject to $a_i^T x_c + r ||a_i||_2 \leq b_i, \quad i = 1, \dots, m$

Quadratic program (QP)

minimize
$$(1/2)x^TPx + q^Tx + r$$
 subject to $Gx \leq h$ $Ax = b$

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



Examples

1/2 |A x - b|_2 = 1/2 (A x - b)' (A x - b) = 1/2 x' A' A x - b' A x + constant

least-squares

minimize 1/2
$$||Ax - b||_2^2$$

Making gradient = 0 A' A x^* - b A = 0 x^* = (A' A)^-1 A b

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

linear program with random cost

minimize
$$\bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E} \, c^T x + \gamma \, \mathbf{var}(c^T x)$$
 subject to $Gx \leq h$, $Ax = b$

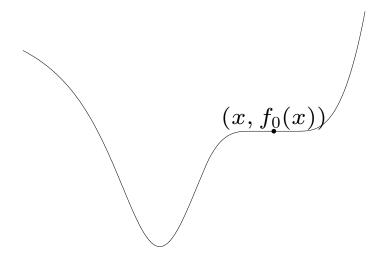
- ullet c is random vector with mean \bar{c} and covariance Σ
- ullet hence, c^Tx is random variable with mean \bar{c}^Tx and variance $x^T\Sigma x$
- \bullet $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)

Quasiconvex optimization

minimize
$$f_0(x)$$
 subject to $f_i(x) \leq 0, \quad i=1,\ldots,m$ $Ax=b$

with $f_0: \mathbf{R}^n \to \mathbf{R}$ quasiconvex, f_1 , . . . , 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.,

$$f_0(x) \le t \iff \phi_t(x) \le 0$$

example

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

with p convex, q concave, and $p(x) \geq 0$, q(x) > 0 on $\operatorname{dom} f_0$ can take $\phi_t(x) = p(x) - tq(x)$:

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

quasiconvex optimization via convex feasibility problems

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

- ullet for fixed t, a convex feasibility problem in x
- ullet 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 $u-l \leq \epsilon$.

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

Linear-fractional program

minimize
$$f_0(x)$$

subject to $Gx \leq h$
 $Ax = b$

linear-fractional program

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

• a quasiconvex optimization problem; can be solved by bisection

Second-order cone programming

minimize
$$f^Tx$$
 subject to $||A_ix + b_i||_2 \le c_i^Tx + d_i, \quad i = 1, \dots, m$ $Fx = g$

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

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

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

(example)

Robust linear programming

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

minimize
$$c^T x$$

subject to $a_i^T x \leq b_i, \quad i = 1, \dots, m,$

there can be uncertainty in c, a_i , b_i

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

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

minimize
$$c^T x$$
 subject to $a_i^T x \leq b_i$ for all $a_i \in \mathcal{E}_i$, $i = 1, \ldots, m$,

ullet stochastic model: a_i is random variable; constraints must hold with probability η

minimize
$$c^T x$$

subject to $\mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \dots, m$

deterministic approach via SOCP

• choose an ellipsoid as \mathcal{E}_i :

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

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

robust LP

minimize
$$c^T x$$

subject to $a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \dots, m$

is equivalent to the SOCP

minimize
$$c^T x$$

subject to $\bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \quad i = 1, \dots, m$

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

a_i'x is constant wrt u, only analyze (P_i u)'x $\sup_{|u|_2 \le 1} u' P_i' x = |P_i' x|_2$ by norm duality

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

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

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

• robust LP

minimize
$$c^T x$$

subject to $\mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \dots, m,$

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

minimize
$$c^Tx$$
 subject to $\bar{a}_i^Tx + \Phi^{-1}(\eta) \|\Sigma_i^{1/2}x\|_2 \leq b_i, \quad i=1,\ldots,m$

Geometric programming

monomial function

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

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

posynomial function: sum of monomials

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

geometric program (GP)

minimize
$$f_0(x)$$

subject to $f_i(x) \leq 1, \quad i = 1, \dots, m$
 $h_i(x) = 1, \quad i = 1, \dots, p$

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 $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ transforms to

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

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

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

geometric program transforms to convex problem

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

$$Gy+d=0$$

Generalized inequality constraints

convex problem with generalized inequality constraints

minimize
$$f_0(x)$$
 subject to $f_i(x) \preceq_{K_i} 0$, $i = 1, \ldots, m$ $Ax = b$

- $f_0: \mathbf{R}^n \to \mathbf{R}$ 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

minimize
$$c^T x$$

subject to $Fx + g \leq_K 0$
 $Ax = b$

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

Semidefinite program (SDP)

minimize
$$c^Tx$$
 subject to $x_1F_1+x_2F_2+\cdots+x_nF_n+G\preceq 0$ $Ax=b$

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

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

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

is equivalent to single LMI

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

(example)

Eigenvalue minimization

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

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

equivalent SDP

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

$$\lambda_{\max}(A) \le t \iff A \le tI$$

Matrix norm minimization

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

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

Assume this is X and you will see