Convex Optimization — Boyd & Vandenberghe

# 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  $f_i(x) \le 0$ ,  $i = 1, ..., m$   
 $h_i(x) = 0$ ,  $i = 1, ..., p$ 

- $x \in \mathbf{R}^n$  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)  $st. x \le 2$ -x \le -3
- $p^{\star} = -\infty$  if problem is unbounded below min x st. x  $\leq$  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

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

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

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

## **Implicit constraints**

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

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

- $\bullet\,$  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  $f_i(x) \le 0, \quad i = 1, \dots, m$   
 $h_i(x) = 0, \quad i = 1, \dots, p$ 

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

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

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

## **Convex optimization problem**

#### standard form convex optimization problem

minimize 
$$f_0(x)$$
  
subject to  $f_i(x) \le 0$ ,  $i = 1, ..., m$   
 $a_i^T x = b_i$ ,  $i = 1, ..., p$ 

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

often written as

minimize 
$$f_0(x)$$
  
subject to  $f_i(x) \le 0$ ,  $i = 1, ..., m$   
 $Ax = b$ 

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

#### example

$$\begin{array}{ll} \mbox{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \mbox{subject to} & f_1(x) = x_1/(1+x_2^2) \leq 0 & \mbox{denominator} > 0 \mbox{ then } \texttt{x1} \leq 0 \\ & h_1(x) = (x_1+x_2)^2 = 0 & \mbox{x1} = -\texttt{x2} \end{array}$$

•  $f_0$  is convex; feasible set  $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$  is convex

- not a convex problem (according to our definition):  $f_1$  is not convex,  $h_1$  is not affine
- equivalent to the convex problem

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

## 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  $f_0(y) < f_0(x)$  (i.e., x not globally optimal) ot in the ocal region

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

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



consider 
$$z = \theta y + (1 - \theta)x$$
 with  $\theta = R/(2||y - x||_2)$   
not in the then  $\theta$  |y-x|\_2 = R/2

local reg

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

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

• 
$$||z - x||_2 = R/2$$
 and  
 $z - x = \theta y + (1 - \theta)x - x$   
 $= \theta (y - x)$   
 $|z - x|_2 = \theta |y - x|_2$   
 $= R/2$   
fo(z) = fo( $\theta y + (1 - \theta)x$ )  
 $\leq \theta fo(y) + (1 - \theta) fo(x)$   
 $= fo(x)$   
fo(z) = fo( $\theta y + (1 - \theta)x$ )  
 $\leq \theta fo(y) + (1 - \theta) fo(x)$   
 $\dots since fo(y) < fo(x)$   
 $= fo(x)$ 

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

## **Optimality criterion for differentiable** $f_0$



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  $f_i(x) \le 0$ ,  $i = 1, ..., m$   
 $Ax = b$ 

is equivalent to

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

where F and  $x_0$  are such that

$$Ax = b \iff x = Fz + x_0$$
 for some  $z$   
A xo = b, A F = 0, then A x = A (F z + x0) = A F z + A x0 = A x0 = b

• introducing equality constraints

minimize 
$$f_0(A_0x + b_0)$$
  
subject to  $f_i(A_ix + b_i) \le 0$ ,  $i = 1, ..., m$ 

is equivalent to

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

• 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

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

• epigraph form: standard form convex problem is equivalent to

minimize (over 
$$x, t$$
)  $t$   
subject to  
 $f_0(x) - t \le 0$   
 $f_i(x) \le 0, \quad i = 1, \dots, m$   
 $Ax = b$ 

• minimizing over some variables

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

is equivalent to

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

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

- 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,

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

piecewise-linear minimization

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

equivalent to an LP

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

Convex optimization problems

#### 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

• hence,  $x_c$ , r can be determined by solving the LP

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



## 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**

least-squares

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

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

linear program with random cost

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

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

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

### **Quasiconvex optimization**

minimize 
$$f_0(x)$$
  
subject to  $f_i(x) \le 0$ ,  $i = 1, ..., m$   
 $Ax = b$ 

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

 $(x, f_0(x))$ 

#### convex representation of sublevel sets of $f_0$

if  $f_0$  is quasiconvex, there exists a family of functions  $\phi_t$  such that:

- $\phi_t(x)$  is convex in x for fixed t
- *t*-sublevel set of  $f_0$  is 0-sublevel set of  $\phi_t$ , *i.e.*,

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

#### example

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

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

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

• for  $t \ge 0$ ,  $\phi_t$  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)**  $\le 0$ 

Convex optimization problems

#### 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)

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

Bisection method for quasiconvex optimization

```
given l \leq p^*, u \geq p^*, 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

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

linear-fractional program

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

• a quasiconvex optimization problem; can be solved by bisection

## Second-order cone programming

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

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

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

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

### **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)

• 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  $a_i \in \mathcal{E}_i$ ,  $i = 1, \dots, m_i$ 

- stochastic model:  $a_i$  is random variable; constraints must hold with probability  $\eta$ 

minimize 
$$c^T x$$
  
subject to  $\operatorname{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \dots, m$ 

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

(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

Convex optimization problems

#### stochastic approach via SOCP

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

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

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

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

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

## **Geometric programming**

#### monomial function

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

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

posynomial function: sum of monomials

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

## geometric program (GP)

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

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 
$$f(x) = cx_1^{a_1} \cdots x_n^{a_n}$$
 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  $\log \left( \sum_{k=1}^{K} \exp(a_{ik}^T y + b_{ik}) \right) \le 0, \quad i = 1, \dots, m$   
 $Gy + d = 0$ 

## Generalized inequality constraints

convex problem with generalized inequality constraints

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

- $f_0: \mathbf{R}^n \to \mathbf{R}$  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  $Fx + g \preceq_K 0$   
 $Ax = b$ 

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

## Semidefinite program (SDP)

minimize 
$$c^T x$$
  
subject to  $x_1F_1 + x_2F_2 + \dots + x_nF_n + G \leq 0$   
 $Ax = b$ 

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

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

### **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

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

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

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

## Matrix norm minimization

minimize 
$$||A(x)||_2 = (\lambda_{\max}(A(x)^T A(x)))^{1/2}$$
  
where  $A(x) = A_0 + x_1 A_1 + \dots + x_n A_n$  (with given  $A_i \in \mathbb{R}^{p \times q}$ )  
equivalent SDP

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

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

Let X = [W B; B'C] Schur complement: D = C - B' W^-1 B If W in S\_++ then X in S\_+ if and only if D in S\_+ Assume this is X and you will see Convex optimization problems