Lecture 7: Convex Optimization Problems

goals:

- define class of convex optimization problems
- equivalent optimization problems
- review common examples

Standard form optimization problem

$$
\begin{array}{lll}
\text{minimize} & f(x) & \\
\text{subject to} & g_i(x) \leq 0 & i = 1, \dots, m \\
 & h_i(x) = 0 & i = 1, \dots, p
\end{array}
$$

We call the problem convex if

- f and g_i are all convex functions
- h_i are affine functions: $h_i(x) = a_i^T x + b_i$

often written: $Ax + b = 0$

(or $Ax = b$)

- The feasible set of a convex optimization problem is a convex set:

  - the feasible set associated w/ each inequality constraint is a sublevel set of a convex function, therefore convex
  - each affine equality constraint defines a (convex) hyperplane
  - their intersection is a convex set

$\boxed{\text{Some Fine Print}}$

- Note however that an optimization problem with convex objective and convex feasible set is not necessarily a convex problem (per our definition)

① $\underline{\underline{Ex}}$

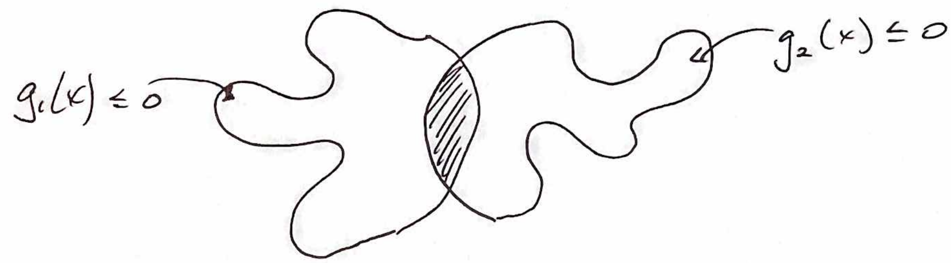
$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & g_1(x) = \frac{x_1}{1+x_2^2} \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0 \end{aligned}$$

- <u>not</u> convex (by our definition): g_1 not convex, h_1 not affine

- equivalent (but not same as) convex problem

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

② In general, feasible set can be convex even when constraint functions are not



- ruled out by our definition

③ Some convex sets don't have an efficient finite representation:

Consider the convex cone of copositive matrices

$$K_{cop} = \{ X \in S^n \mid y^T X y \geq 0, \; y \geq 0 \}$$

Its dual cone is the set of completely positive matrices

$$K_{cop}^* = \text{conv}\{ x x^T \mid x \geq 0 \}$$

- It's possible to exactly reformulate non-convex quadratic, mixed integer problems (known to be computationally hard) as problems with convex objectives over these cones!
- Issue is that these cones don't have an efficient finite representation: NP-hard to check if a given matrix is in the cone

** Computational tractability and convexity involve algebraic representation, not just geometry!

A feasibility problem

$$
\begin{array}{lll}
\text{find} & x & \\
\text{subject to} & g_i(x) \le 0 & i = 1, \dots, m \\
& h_i(x) = 0 & i = 1, \dots, p
\end{array}
$$

or any constant

is a special case of standard form with $f(x) = 0$

Equivalent Optimization Problems

An optimization problem is (informally) called equivalent to another if its solution can be easily obtained from the other, and vice versa

- useful for both analysis and algorithms

- some problems that appear non convex, may actually be convex in disguise, with an appropriate transformation to an equivalent problem
    - now an art, lots of research into understanding this more systematically

- Some common transformations that preserve convexity

① eliminating equality constraints

over z

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i=1,\dots,m \\ & Ax = b & \end{array} \iff \begin{array}{ll} \text{minimize} & f(Fz + x_0) \\ \text{subject to} & g_i(Fz + x_0) \le 0 \\ & i=1,\dots,m \end{array}$$

where $Ax = b \iff x = Fz + x_0$ for some z

② introducing equality constraints

over x, y_i

$$\begin{array}{ll} \text{minimize} & f(A_0 x + b_0) \\ \text{subject to} & g_i(A_i x + b_i) \le 0 \\ & i=1,\dots,m \end{array} \iff \begin{array}{lll} \text{minimize} & f(y_0) & \\ \text{subject to} & g_i(y_i) \le 0 & i=0,1,\dots,m \\ & y_i = A_i x_0 + b_i & \end{array}$$

③ introducing slack variables for linear inequalities

over x, s

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & a_i^T x \le b_i \\ & i=1,\dots,m \end{array} \iff \begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & a_i^T x + s_i = b_i & \\ & s_i \ge 0 & i=1,\dots,m \end{array}$$

④ epigraph form

over x, t

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i=1,\dots,m \\ & Ax = b & \end{array} \iff \begin{array}{lll} \text{minimize} & t & \\ \text{subject to} & f(x) \le t & \\ & g_i(x) \le 0 & i=1,\dots,m \\ & Ax = b & \end{array}$$

⑤ partial minimization

$$\begin{array}{ll} \text{minimize} & f(x_1, x_2) \\ \text{subject to} & g_i(x_1) \le 0 \\ & i = 1, \dots, m \end{array} \iff \begin{array}{ll} \text{minimize} & \bar{f}(x_1) \\ \text{subject to} & g_i(x_1) \le 0 \\ & i = 1, \dots, m \end{array}$$

where $\bar{f}(x_1) = \inf_{x_2} f(x_1, x_2)$

⑥ monotone transformations of objective and constraints

- via composition rules

⑦ lifting problems to higher dimensional spaces by redefining variables ; convex relaxations

- sometimes can transform non-convex problems to convex ones exactly (more later)
- often gives a useful convex relaxation of non-convex problems, even when inexact

Ex

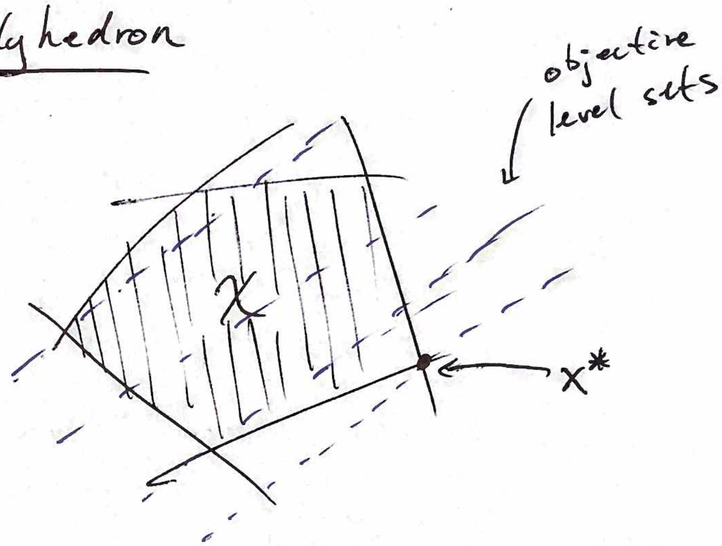
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i \\ & x_i \in \{0, 1\} \end{array} \xrightarrow[\text{relaxation}]{} \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i \\ & x_i \in [0, 1] \end{array}$$

- used extensively in branch-and-bound methods for mixed-integer programs

Common Convex Optimization Problems + Examples

- Linear Program (LP)
  - linear cost + constraint functions
  - feasible set is a <u>polyhedron</u>

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Gx \le h \\ & Ax = b \end{array}$$



- <u>very</u> common in applications
  - variety of planning + scheduling problems
  - problems involving $\|\cdot\|_1$ or $\|\cdot\|_\infty$.
  - piecewise linear problems
  - Nash equilibria in zero-sum games
  - exact formulations of several important combinatorial optimization problems on graphs
    - maximum flow, minimum cut
    - shortest path
    - bipartite matching
  - relaxations of combinatorial problems + use in branch-and-bound methods

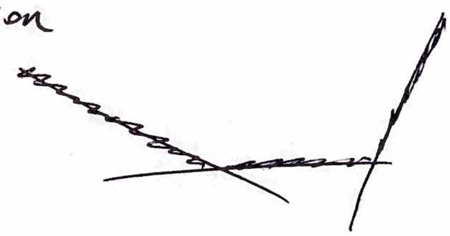
Ex Diet problem: choose quantities $x_1, \dots, x_n$ of n foods

- one unit of food j costs c_j, has amount a_{ij} of nutrient i
- healthy diet requires at least b_i of nutrient i
- finding cheapest healthy diet is an LP:

$$\begin{array}{lll} \text{minimize} & c^T x & \text{(total cost)} \\ \text{subject to} & Ax \geq b & \text{(health constraint)} \\ & x \geq 0 & \text{(nonnegative amounts)} \end{array}$$

Ex Piecewise Linear Minimization

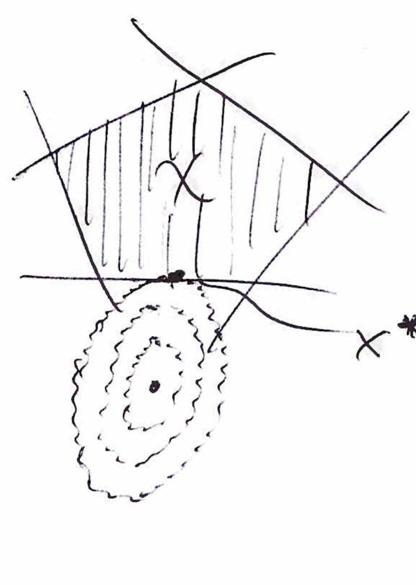
$$\text{minimize} \quad \max_i \{a_i^T x + b_i\}$$



$$\Updownarrow$$

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

- Quadratic Program (QP)
  - quadratic cost, linear constraints
  - convex if $P \succeq 0$



$$\begin{array}{ll} \text{minimize} & \frac{1}{2} x^T P x + q^T x \\ \text{subject to} & Gx \leq h \\ & Ax = b \end{array}$$

- least squares is canonical example

Ex LP with random cost

- cost parameter $c \in \mathbb{R}^n$ is a random vector with mean $\bar{c}$ and covariance matrix $\Sigma \in S_+^n$
- penalize cost variations as measured by Σ

$$
\begin{array}{ll}
\text{minimize} & E[c^T x] + \gamma \operatorname{Var}[c^T x] = \bar{c}^T x + \gamma x^T \Sigma x \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

- See also Markowitz portfolio optimization

<u>Quadratically Constrained Quadratic Program</u> (QCQP)

$$
\begin{array}{lll}
\text{minimize} & \frac{1}{2} x^T P x + q^T x & \\
\text{subject to} & \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0 & i = 1, \dots, m
\end{array}
$$

$$P, P_i \in S^n, \quad q, q_i \in \mathbb{R}^n, \quad r_i \in \mathbb{R}$$

- quadratic costs and constraints
- feasible set is an intersection of ellipsoids (if problem is convex)
- convex if $P, P_i \succeq 0$

- Second-Order Cone Program (SOCP)

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

$$A_i \in \mathbb{R}^{n_i \times n}, \quad F \in \mathbb{R}^{p \times n}$$

- inequalities are called second-order cone constraints

$$(A_i x + b_i, \ c_i^T x + d) \in K_{soc}^{n+1} = \{(z,t) \in \mathbb{R}^{n+1} \mid \|z\|_2 \le t\}$$

- strictly generalizes QCQP

Ex Robust LP

Consider the LP:
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i \quad i=1,\dots,m \end{array}$$

where there is uncertainty in parameters $a_i \in \mathbb{R}^n$.

Assume a_i is only known to lie in an ellipsoid

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

Robust LP: require constraint satisfaction for <u>all</u> possible values of a_i

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

But $a_i^T x \le b_i \ \ \forall a_i \in \mathcal{E}_i \iff \sup\{a_i^T x \mid a_i \in \mathcal{E}_i\} \le b_i$

The LHS can be written

$$\sup\{a_i^T x \mid a_i \in \mathcal{E}_i\} = \bar{a}_i^T x + \sup\{u^T P_i^T x \mid \|u\|_2 \le 1\}$$
$$= \bar{a}_i^T x + \|P_i^T x\|_2$$

Thus the robust LP is equivalent to the SOCP

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

- many other examples; see e.g. "Applications of second order cone programming" by Lobo et al (1998)

Generalized standard form using generalized (cone) inequalities

---

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \preceq_{K_i} 0 & i=1,\dots,m \\ & Ax = b & \end{array}$$

where K_i are proper cones

- already seen special cases:
  - $K_i = \mathbb{R}_{++}$: LP
  - $K_i = K_{soc}$: SOCP
- another important special case: $K_i = S_+^n$: Semidefinite Programming (SDP)

- Semidefinite Program (SDP)

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

Variable: $x \in \mathbb{R}^n$, data: $G, F_1, \dots, F_n \in S^k$, $A \in \mathbb{R}^{p \times n}$
$c \in \mathbb{R}^n$

- feasible set is intersection of an inverse image of S_+^n under affine function $\sum_i x_i F_i + G$ with an affine set
  - sometimes called a spectrahedron
  - far richer geometric objects than polyhedra (can have curved surfaces)

- other forms:

① "Standard" form

$$\begin{array}{ll} \text{minimize} & \operatorname{trace}(CX) \\ \text{subject to} & \operatorname{trace}(A_i X) = b_i \quad i = 1, \dots, m \\ & X \succeq 0 \end{array}$$

variable: $X \in S^n$,
data: $C, A_1, \dots, A_m \in S^n$

② "Inequality form"

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

variables: $x \in \mathbb{R}^n$
data: $A_1, \dots, A_n, B \in S^k$
$c \in \mathbb{R}^n$

- Why SDP?
  - natural generalization of LP, but with much richer expressive power
    - in fact, we have

$$\text{LP} \subseteq \text{QP} \subseteq \text{QCQP} \subseteq \text{SOCP} \subseteq \text{SDP}$$

  - convex problems, with efficient software
  - occurs often in many systems and control design and analysis problems (much more later)
    - (Lyapunov) stability analysis
    - optimal control design
    - invariant set computations
    - system input/output properties
    - etc.
  - nice theory, much of it mirroring LP, and connections to many other areas of mathematics

$\underline{\underline{\text{Ex}}}$ Eigenvalue minimization

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

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n \quad (A_i \in S^k)$

Write in epigraph form:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \lambda_{max}(A(x)) \leq t \end{array}$$

Note that $\lambda_{max}(A(x)) \leq t \iff A(x) \preceq tI$

Equivalent SDP:

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

Quasiconvex Optimization

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \qquad i=1,\dots,m \\ & Ax = b \end{aligned}$$

with f quasiconvex and g_i convex

- convex representation of sublevel sets of f:

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

① $\phi_t(x)$ is convex in x for fixed t

② t-sublevel set of f is 0-sublevel set of ϕ_t:

$$f(x) \leq t \iff \phi_t(x) \leq 0$$

- Ex $\quad f(x) = \dfrac{p(x)}{q(x)}$ with p convex, $p(x) \geq 0$; q concave, $q(x) > 0$ on $\mathrm{dom}(f)$

Consider $\phi_t(x) = p(x) - t\,q(x)$

Clearly:

- $\phi_t(x)$ convex in x (for fixed $t \geq 0$)
- $\dfrac{p(x)}{q(x)} \leq t \iff \phi_t(x) \leq 0$

$$
\begin{array}{ll}
\text{minimize} & f(x) \\
\text{subject to} & g_i(x) \le 0 \\
& Ax = b
\end{array}
\quad \underset{\substack{\text{epigraph} \\ \text{form}}}{\Longleftrightarrow} \quad
\overset{(*)}{
\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \phi_t \le 0 \\
& g_i(x) \le 0 \\
& Ax = b
\end{array}}
$$

- for fixed t, $\overset{(*)}{\wedge}$ a convex feasibility problem in x
  - feasible $\Longrightarrow t \ge f^*$
  - infeasible $\Longrightarrow t \le f^*$

  $\Big\}$ use bisection!

Bisection method for quasiconvex optimization

---

given $f^* \in [l, u]$, tolerance ε

while $u - l > \varepsilon$

① $t = \frac{1}{2}(l + u)$

② solve convex feasibility problem $(*)$ for fixed t

③ if $(*)$ feasible, set $u = t$

else, set $l = t$

- requires exactly $\log_2\left(\frac{u-l}{\varepsilon}\right)$ iterations
  - e.g., $\frac{u-l}{\varepsilon} = 10^{12} \Longrightarrow 40$ iterations