EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 8

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

April 26, 2012

Caltech

Outline

Standard Form of Optimization Problems

Most of the optimization problems we will be concerned with will be assumed to be expressible in the following standardized form.

Standard Form of General Optimization Problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array} .$$

- $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable.
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective or cost function.
- $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \ldots, m$, are the inequality constraint functions.
- $h_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \ldots, p$, are the equality constraint functions.
- $\mathcal{D} \triangleq \bigcap_{k=0}^{m} \operatorname{dom}(f_k) \cap \bigcap_{k=1}^{p} \operatorname{dom}(h_k)$ is the domain of the problem.

Optimal Value:

$$p^\star = \inf\{f_0(\mathbf{x}) : f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m\,,\ h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p\} .$$

- $p^\star = \infty$ if the problem is *infeasible* (meaning no $\mathbf{x}$ satisfies the constraints).
- $p^\star = -\infty$ if the problem is *unbounded below*.

Global and Local Optimality

Feasibility: A point $\mathbf{x}$ is said to be *feasible* if $\mathbf{x} \in \mathrm{dom}(f_0)$ and it satisfies the constraints.

(Globally) Optimal Points: A feasible point $\mathbf{x}^\star$ is said to be *(globally) optimal* if $f_0(\mathbf{x}^\star) = p^\star$. The set of all optimal points is the *optimal set* and is given by

$$\mathcal{X}_{\mathrm{opt}} \triangleq \{\mathbf{x} : f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m\,,\ h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p\,,\ f_0(\mathbf{x}) = p^\star\}\ .$$

If $\mathcal{X}_{\mathrm{opt}}$ is nonempty, we say that the optimal value is *achieved* and the problem is *solvable*. Otherwise, the optimal value is not attained or not achieved.

A feasible point $\mathbf{x}$ with $f_0(\mathbf{x}) \leq p^\star + \epsilon$ (where $\epsilon > 0$) is said to be ϵ*-suboptimal*, and the set of all ϵ-suboptimal points is called the ϵ*-suboptimal set* for the problem.

Locally Optimal Points: A feasible point $\mathbf{x}_{\mathrm{lo}}$ is said to be *locally optimal* if there exists an $R > 0$ such that $\mathbf{x}_{\mathrm{lo}}$ is optimal for the problem:

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{z}) \\ \text{subject to} & f_k(\mathbf{z}) \leq 0\,,\ k = 1, \ldots, m\,,\ h_k(\mathbf{z}) = 0\,,\ k = 1, \ldots, p\,, \quad . \\ & ||\mathbf{z} - \mathbf{x}_{\mathrm{lo}}||_2 \leq R \end{array}$$

Examples: (with $n = 1$, $m = p = 0$)

- $f_0(x) = 1/x$, $\mathrm{dom}(f_0) = \mathbb{R}_{++}$: $p^\star = 0$, no optimal point.
- $f_0(x) = -\log x$, $\mathrm{dom}(f_0) = \mathbb{R}_{++}$: $p^\star = -\infty$, no optimal point.
- $f_0(x) = x \log x$, $\mathrm{dom}(f_0) = \mathbb{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal.
- $f_0(x) = x^3 - 3x$, $\mathrm{dom}(f_0) = \mathbb{R}$: $p^\star = -\infty$, no optimal point, local optimum at $x = 1$.

The Feasibility Problem

The *feasibility problem* is the problem of determining whether or not the feasible set is empty. It is given by

$$\begin{array}{ll} \text{find} & \mathbf{x} \\ \text{subject to} & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array} .$$

The feasibility problem can be considered a special case of the general problem with $f_0(\mathbf{x}) = 0$, namely

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array} .$$

- We have $p^\star = 0$ if the constraints are feasible; any feasible $\mathbf{x}$ is optimal in this case.
- We have $p^\star = \infty$ if the constraints are infeasible.

Change of Variables and Function Transformations

Two optimization problems will be called *equivalent* if, from a solution of one problem, a solution of the other is readily found, and vice versa. An appropriate change of variables or a transformation of the objective and constraint functions yield equivalent problems.

Change of Variables:
Suppose $\boldsymbol{\phi} : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one, with $\boldsymbol{\phi}(\text{dom}(\boldsymbol{\phi})) \supseteq \mathcal{D}$. With the *change of variable* $\mathbf{x} = \boldsymbol{\phi}(\mathbf{z})$, if we define the functions

$$\widetilde{f}_k(\mathbf{z}) \triangleq f_k(\boldsymbol{\phi}(\mathbf{z}))\ ,\ k = 0, \ldots, m\,,\ \widetilde{h}_k(\mathbf{z}) \triangleq h_k(\boldsymbol{\phi}(\mathbf{z}))\ ,\ k = 1, \ldots, p\,,$$

then the following problem is equivalent to the original standard form problem:

$$\begin{array}{ll} \text{minimize} & \widetilde{f}_0(\mathbf{z}) \\ \text{subject to} & \widetilde{f}_k(\mathbf{z}) \leq 0\,,\ k = 1, \ldots, m\,,\ \widetilde{h}_k(\mathbf{z}) = 0\,,\ k = 1, \ldots, p \end{array}.$$

Transformation of Objective and Constraint Functions:
Suppose $\psi_0 : \mathbb{R} \to \mathbb{R}$ is monotone increasing, $\psi_1, \ldots, \psi_m : \mathbb{R} \to \mathbb{R}$ satisfy $\phi_k(u) \leq 0$ if and only if $u \leq 0$, and $\psi_{m+1}, \ldots, \psi_{m+p} : \mathbb{R} \to \mathbb{R}$ satisfy $\phi_k(u) = 0$ if and only if $u = 0$. If we define the functions

$$\widetilde{f}_k(\mathbf{x}) \triangleq \psi_k(f_k(\mathbf{x}))\ ,\ k = 0, \ldots, m\,,\ \widetilde{h}_k(\mathbf{x}) \triangleq \psi_{m+k}(h_k(\mathbf{x}))\ ,\ k = 1, \ldots, p\,,$$

then the following problem is equivalent to the original standard form problem:

$$\begin{array}{ll} \text{minimize} & \widetilde{f}_0(\mathbf{x}) \\ \text{subject to} & \widetilde{f}_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m\,,\ \widetilde{h}_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array}.$$

Eliminating/Introducing Equality Constraints

Eliminating Equality Constraints:
Suppose that there is a function $\boldsymbol{\phi} : \mathbb{R}^k \to \mathbb{R}^n$ such that $\mathbf{x}$ satisfies $h_k(\mathbf{x}) = 0$ for $k = 1, \ldots, p$, if and only if there is some $\mathbf{z} \in \mathbb{R}^k$ such that $\mathbf{x} = \boldsymbol{\phi}(\mathbf{z})$. Then, the optimization problem

$$\begin{array}{ll} \text{minimize} & \widetilde{f}_0(\mathbf{z}) = f_0(\boldsymbol{\phi}(\mathbf{z})) \\ \text{subject to} & \widetilde{f}_k(\mathbf{z}) = f_k(\boldsymbol{\phi}(\mathbf{z})) \leq 0\,,\ k = 1, \ldots, m \end{array},$$

is equivalent to the original standard form problem. The transformed problem has variable $\mathbf{z} \in \mathbb{R}^k$, m inequality constraints, and no equality constraints.

Introducing Equality Constraints:
In many cases, it will actually be advantageous to *introduce* equality constraints and new variables into a problem. As an example, consider the problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{A}_0\mathbf{x} + \mathbf{b}_0) \\ \text{subject to} & f_k(\mathbf{A}_k\mathbf{x} + \mathbf{b}_k) \leq 0\,,\ k = 1, \ldots, m\,,\ h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array},$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{A}_k \in \mathbb{R}^{\ell_k \times n}$, $\mathbf{b}_k \in \mathbb{R}^{\ell_k}$, and $f_k : \mathbb{R}^{\ell_k} \to \mathbb{R}$. Introducing new variables $\mathbf{y}_k \in \mathbb{R}^{\ell_k}$ and new equality constraints $\mathbf{y}_k = \mathbf{A}_k\mathbf{x} + \mathbf{b}_k$ for $k = 0, \ldots, m$ yields the equivalent problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{y}_0) \\ \text{subject to} & f_k(\mathbf{y}_k) \leq 0\,,\ k = 1, \ldots, m\,,\ h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \\ & \mathbf{y}_k = \mathbf{A}_k\mathbf{x} + \mathbf{b}_k\,,\ k = 0, \ldots, m \end{array}.$$

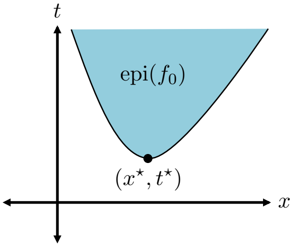
This problem has $\ell_0 + \cdots + \ell_m$ new variables and $\ell_0 + \cdots + \ell_m$ equality constraints.

Epigraph Problem Form

The *epigraph form* of the standard form problem is the problem

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_0(\mathbf{x}) - t \leq 0 \\ & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & h_k(\mathbf{x}) = 0\,,\ k = 1, \ldots, p \end{array},$$

with variables $\mathbf{x} \in \mathbb{R}^n$ and $t \in \mathbb{R}$.



It is easy to see that it is equivalent to the original problem; $(\mathbf{x}, t)$ is optimal for the epigraph form if and only if $\mathbf{x}$ is optimal for the original problem and $t = f_0(\mathbf{x})$. Note that the objective function for the epigraph form problem is a *linear* function of $(\mathbf{x}, t)$.

Standard Form Convex Optimization Problem

In an abstract sense, a convex optimization problem is minimizing a convex function over a convex set. Here, we will call a problem a convex optimization problem if it can be expressed in the following way.

Standard Form of Convex Optimization Problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_k(\mathbf{x}) \le 0\,,\ k = 1, \ldots, m \\ & \mathbf{a}_k^T \mathbf{x} = b_k\,,\ k = 1, \ldots, p \end{array} \equiv \begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_k(\mathbf{x}) \le 0\,,\ k = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}.$$

- The objective function f_0 and inequality constraint functions $f_1, \ldots, f_m$ are *convex*.
- The equality constraints, characterized by $\mathbf{a}_k \in \mathbb{R}^n$ and $b_k \in \mathbb{R}$ for $k = 1, \ldots, p$, or $\mathbf{A} \in \mathbb{R}^{p \times n}$ and $\mathbf{b} \in \mathbb{R}^p$, are *affine*.

Important Property:

- The feasible set of a convex optimization problem is *convex*.

Local and Global Optimality

Locally Optimal Points are Globally Optimal:

Any locally optimal point of a convex optimization problem is also (globally) optimal.

Proof: (By Contradiction)

Suppose $\mathbf{x}$ is locally optimal and $\mathbf{y}$ is optimal with $f_0(\mathbf{y}) < f_0(\mathbf{x})$. As $\mathbf{x}$ is locally optimal, this means that there is an $R > 0$ such that

$$\mathbf{z} \text{ is feasible}\,, \ \left|\left|\mathbf{z} - \mathbf{x}\right|\right|_2 \leq R \Longrightarrow f_0(\mathbf{z}) \geq f_0(\mathbf{x}) \ .$$

Consider $\mathbf{z} = (1-\theta)\,\mathbf{x} + \theta\mathbf{y}$ with $\theta = R/\left(2\left|\left|\mathbf{y}-\mathbf{x}\right|\right|_2\right)$. Then, we have the following.

- As $\left|\left|\mathbf{y}-\mathbf{x}\right|\right|_2 > R$, we have $0 < \theta < 1/2$.
- Since $\mathbf{z}$ is a convex combination of two feasible points $\mathbf{x}$ and $\mathbf{y}$, it is also feasible.
- We have $\left|\left|\mathbf{z}-\mathbf{x}\right|\right|_2 = R/2 < R$.
- By the convexity of f_0, we have

  $$f_0(\mathbf{z}) \leq (1-\theta)\, f_0(\mathbf{x}) + \theta f_0(\mathbf{y}) < f_0(\mathbf{x}) \ ,$$

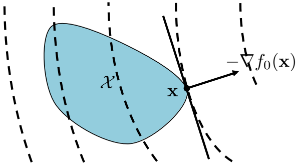
  which contradicts our assumption that $\mathbf{x}$ is locally optimal.

Optimality Criterion for Differentiable Objectives

A point $\mathbf{x}$ is optimal if and only if it is feasible and

$$(\nabla f_0(\mathbf{x}))^T (\mathbf{y} - \mathbf{x}) \geq 0 \,,$$

for all feasible $\mathbf{y}$. If nonzero, $-\nabla f_0(\mathbf{x})$ defines a supporting hyperplane to the feasible set $\mathcal{X}$ at $\mathbf{x}$.



- **Unconstrained problem:** minimize $f_0(\mathbf{x})$.
  A point $\mathbf{x}$ is optimal if and only if

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

- **Equality constrained problem:** minimize $f_0(\mathbf{x})$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$.
  A point $\mathbf{x}$ is optimal if and only if there exists a $\boldsymbol{\nu}$ such that

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

- **Minimization over nonnegative orthant:** minimize $f_0(\mathbf{x})$ subject to $\mathbf{x} \succeq \mathbf{0}$.
  A point $\mathbf{x}$ is optimal if and only if

$$\mathbf{x} \in \mathrm{dom}(f_0) \ , \ \mathbf{x} \succeq \mathbf{0} \,, \ \begin{cases} [\nabla f_0(\mathbf{x})]_k \geq 0 \,, & x_k = 0 \\ [\nabla f_0(\mathbf{x})]_k = 0 \,, & x_k > 0 \end{cases} \,.$$

Equivalent Convex Problems

- **Eliminating equality constraints:**

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_k(\mathbf{x}) \le 0\,,\ k = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } \mathbf{z}) & f_0(\mathbf{F}\mathbf{z} + \mathbf{x}_0) \\ \text{subject to} & f_k(\mathbf{F}\mathbf{z} + \mathbf{x}_0) \le 0\,,\ k = 1, \ldots, m \end{array},$$

where $\mathbf{F}$ and $\mathbf{x}_0$ are such that

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Longleftrightarrow \mathbf{x} = \mathbf{F}\mathbf{z} + \mathbf{x}_0\,.$$

In other words, $\mathcal{R}(\mathbf{F}) = \mathcal{N}(\mathbf{A})$ and $\mathbf{x}_0$ is a particular solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$.

- **Introducing equality constraints:**

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{A}_0\mathbf{x} + \mathbf{b}_0) \\ \text{subject to} & f_k(\mathbf{A}_k\mathbf{x} + \mathbf{b}_k) \le 0\,,\ k = 1, \ldots, m \end{array},$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } \mathbf{x}, \mathbf{y}_k) & f_0(\mathbf{y}_0) \\ \text{subject to} & f_k(\mathbf{y}_k) \le 0\,,\ k = 1, \ldots, m \\ & \mathbf{y}_k = \mathbf{A}_k\mathbf{x} + \mathbf{b}_k\,,\ k = 0, 1, \ldots, m \end{array}.$$

Equivalent Convex Problems (Continued)

- **Introducing slack variables for linear inequalities:**

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x} \leq b_k \,,\ k = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } \mathbf{x}, \mathbf{s}) & f_0(\mathbf{x}) \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x} + s_k = b_k \,,\ k = 1, \ldots, m \,,\ s_k \geq 0 \,,\ k = 1, \ldots, m \end{array} .$$

- **Epigraph form:** The standard form convex problem is equivalent to

$$\begin{array}{ll} \text{minimize (over } \mathbf{x}, t) & t \\ \text{subject to} & f_0(\mathbf{x}) - t \leq 0 \,,\ f_k(\mathbf{x}) \leq 0 \,,\ k = 1, \ldots, m \,,\ \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

- **Minimizing over some variables:**

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}_1, \mathbf{x}_2) \\ \text{subject to} & f_k(\mathbf{x}_1) \leq 0 \,,\ k = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \widetilde{f}_0(\mathbf{x}_1) \\ \text{subject to} & f_k(\mathbf{x}_1) \leq 0 \,,\ k = 1, \ldots, m \end{array} ,$$

where $\widetilde{f}_0(\mathbf{x}_1) = \inf_{\mathbf{x}_2} \{ f_0(\mathbf{x}_1, \mathbf{x}_2) \}$.

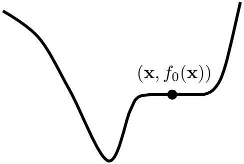
Quasiconvex Optimization Problems

The following is the standard form of a *quasiconvex optimization problem*.

Standard Form of Quasiconvex Optimization Problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_k(\mathbf{x}) \le 0\,,\ k = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}\ .$$

- The objective function $f_0 : \mathbb{R}^n \to \mathbb{R}$ is *quasiconvex*.
- The inequality constraint functions $f_1, \ldots, f_m$ are *convex*.
- The equality constraints, characterized by $\mathbf{a}_k \in \mathbb{R}^n$ and $b_k \in \mathbb{R}$ for $k = 1, \ldots, p$, or $\mathbf{A} \in \mathbb{R}^{p \times n}$ and $\mathbf{b} \in \mathbb{R}^p$, are *affine*.



Problems of this nature can have locally optimal points that are not (globally) optimal.

Convex Representation of Objective Sublevel Sets

If f_0 is quasiconvex, there exists a family of functions ϕ_t such that:

- $\phi_t(\mathbf{x})$ is convex in $\mathbf{x}$ for fixed t,
- the t-sublevel set of f_0 is the 0-sublevel set of ϕ_t, i.e.,

$$f_0(\mathbf{x}) \leq t \Longleftrightarrow \phi_t(\mathbf{x}) \leq 0\,.$$

Example:

Suppose f_0 is given by

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

where p is convex and q is concave with $p(\mathbf{x}) \geq 0$ and $q(\mathbf{x}) > 0$ on $\mathrm{dom}(f_0)$. Then, if we take

$$\phi_t(\mathbf{x}) = p(\mathbf{x}) - tq(\mathbf{x})\ ,$$

we have the following:

- for $t \geq 0$, ϕ_t is convex in $\mathbf{x}$,
- $p(\mathbf{x})\,/q(\mathbf{x}) \leq t$ if and only if $\phi_t(\mathbf{x}) \leq 0$.

Bisection Method for Quasiconvex Problems

Using a convex representation of the sublevel sets of a quasiconvex objective, we can solve a quasiconvex optimization problem via a series of convex feasibility problems. Each of these feasibility problems is of the following form.

$$\begin{array}{ll} \text{find} & \mathbf{x} \\ \text{subject to} & \phi_t(\mathbf{x}) \leq 0 \\ & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} . \tag{1}$$

- For fixed t, this is a convex feasibility problem in $\mathbf{x}$.
- If feasible, we can conclude that $t \geq p^\star$; if infeasible, then $t \leq p^\star$.

This leads to the following *bisection method* for solving quasiconvex optimization problems.

Bisection Method for Solving Quasiconvex Optimization Problems:

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, then $u := t$; **else** $l := t$.

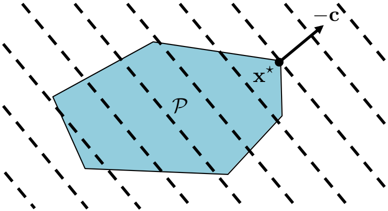
Until $u - l \leq \epsilon$.

Linear Programs

When the objective and constraint functions are all affine, the problem is called a *linear program* (LP). A general LP has the following form:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} + d \\ \text{subject to} & \mathbf{G}\mathbf{x} \preceq \mathbf{h} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array},$$

where $\mathbf{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$, $\mathbf{G} \in \mathbb{R}^{m \times n}$, $\mathbf{h} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, and $\mathbf{b} \in \mathbb{R}^p$.



- The feasible set is always a polyhedron.
- One of the vertices of the polyhedron is always an optimal point.

Examples of LPs

- **Piecewise-linear minimization:**

$$\text{minimize} \quad \max_{k=1,\ldots,m} \left(\mathbf{a}_k^T \mathbf{x} + b_k \right) \,,$$

is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x} + b_k \leq t \,, \; k = 1, \ldots, m \end{array} .$$

- **Chebyshev inequalities:**
Consider a discrete random variable x on a set $\{u_1, \ldots, u_n\} \subseteq \mathbb{R}$. The *probability mass function* (pmf) is a vector $\mathbf{p} \in \mathbb{R}^n$ with $p_k = \Pr\{x = u_k\}$ and satisfies $\mathbf{p} \succeq \mathbf{0}$ and $\mathbf{1}^T \mathbf{p} = 1$. For any function f, we have

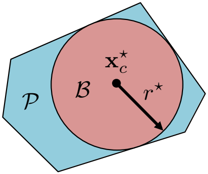
$$E[f(x)] = \sum_{k=1}^{n} p_k f(u_k) = \mathbf{f}^T \mathbf{p} \,, \text{ where } \mathbf{f} \triangleq \left[\begin{array}{ccc} f(u_1) & \cdots & f(u_n) \end{array} \right]^T \,,$$

which is *linear* in $\mathbf{p}$. In the Chebyshev inequality problem, we do not know $\mathbf{p}$, but we are given upper and lower bounds on expected values of some functions of x. From this information, we wish to calculate the upper and lower bounds of the expected value of some target function $E[f_0(x)] = \mathbf{a}_0^T \mathbf{p}$. This leads to the following two LPs.

$$\begin{array}{ll} \text{minimize/maximize} & \mathbf{a}_0^T \mathbf{p} \\ \text{subject to} & \mathbf{p} \succeq \mathbf{0} \,, \; \mathbf{1}^T \mathbf{p} = 1 \,, \; \alpha_k \leq \mathbf{a}_k^T \mathbf{p} \leq \beta_k \,, \; k = 1, \ldots, m \end{array} .$$

Examples of LPs (Continued)

- **Chebyshev center of a polyhedron:**
  The *Chebyshev center* of
  $$\mathcal{P} = \left\{ \mathbf{x} : \mathbf{a}_k^T \mathbf{x} \leq b_k \,,\ k = 1, \ldots, m \right\}$$
  is the center $\mathbf{x}_c$ of the largest inscribed ball (with radius r) given by
  $$\mathcal{B} = \{ \mathbf{x}_c + \mathbf{u} : ||\mathbf{u}|| \leq r \} \ .$$
  - Note that $\mathbf{a}_k^T \mathbf{x} \leq b_k$ for all $\mathbf{x} \in \mathcal{B}$ if and only if
    $$\sup \left\{ \mathbf{a}_k^T (\mathbf{x}_c + \mathbf{u}) : ||\mathbf{u}|| \leq r \right\} = \mathbf{a}_k^T \mathbf{x}_c + r \, ||\mathbf{a}_k||_{\circledast} \leq b_k \ .$$
  - Hence, $\mathbf{x}_c$ and r can be determined by solving the LP
    $$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x}_c + r \, ||\mathbf{a}_k||_{\circledast} \leq b_k \,,\ k = 1, \ldots, m \end{array} \ .$$



Linear-Fractional Programs

The problem of minimizing a ratio of affine functions over a polyhedron is called a *linear-fractional program* and is given by

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & \mathbf{G}\mathbf{x} \preceq \mathbf{h}\,,\ \mathbf{A}\mathbf{x} = \mathbf{b} \end{array},$$

where the objective function is

$$f_0(\mathbf{x}) = \frac{\mathbf{c}^T\mathbf{x} + d}{\mathbf{e}^T\mathbf{x} + f}\,,\ \mathrm{dom}(f_0) = \left\{\mathbf{x} : \mathbf{e}^T\mathbf{x} + f > 0\right\}.$$

- The objective is quasilinear, and so the problem is a quasiconvex optimization problem. It can be solved by bisection.
- With the transformation

  $$\mathbf{y} = \frac{\mathbf{x}}{\mathbf{e}^T\mathbf{x} + f}\,,\ z = \frac{1}{\mathbf{e}^T\mathbf{x} + f}\,,$$

  the problem can be shown to be equivalent to the following LP:

  $$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{y} + dz \\ \text{subject to} & \mathbf{G}\mathbf{y} \preceq \mathbf{h}z\,,\ \mathbf{A}\mathbf{y} = \mathbf{b}z \\ & \mathbf{e}^T\mathbf{y} + fz = 1\,,\ z \geq 0 \end{array}.$$

Generalized Linear-Fractional Programs

A generalization of the linear-fractional program, called the *generalized linear-fractional program*, occurs when the objective is given by

$$f_0(\mathbf{x}) = \max_{k=1,\ldots,r} \frac{\mathbf{c}_k^T\mathbf{x} + d_k}{\mathbf{e}_k^T\mathbf{x} + f_k} \,,\ \mathrm{dom}(f_0) = \left\{\mathbf{x} : \mathbf{e}_k^T\mathbf{x} + f_k > 0 \,,\ k = 1, \ldots, r\right\} .$$

The objective function in this case is quasiconvex, and so the problem is quasiconvex. It can be solved by bisection.

Example: *Von Neumann model of a growing economy*

$$\begin{array}{ll} \text{maximize (over } \mathbf{x}, \mathbf{x}^+\text{)} & \displaystyle\min_{k=1,\ldots,n} \frac{x_k^+}{x_k} \\ \text{subject to} & \mathbf{x}^+ \succeq \mathbf{0} \,,\ \mathbf{B}\mathbf{x}^+ \preceq \mathbf{A}\mathbf{x} \end{array} .$$

- $\mathbf{x}, \mathbf{x}^+ \in \mathbb{R}^n$: activity levels of n sectors of an economy, in the current and next period, respectively.
- $[\mathbf{A}\mathbf{x}]_k \,,\ \left[\mathbf{B}\mathbf{x}^+\right]_k$: amounts of k-th good produced and consumed, respectively.
- x_k^+/x_k: growth rate of k-th sector.

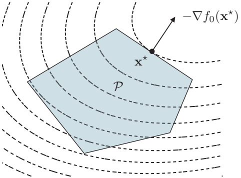
The objective is to allocate activity to maximize the growth rate of the slowest growing sector.

Quadratic Programs

When the objective is convex and quadratic and the constraint functions are all affine, the problem is called a *quadratic program* (QP). A general QP has the following form:

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & \mathbf{G}\mathbf{x} \preceq \mathbf{h}\,,\, \mathbf{A}\mathbf{x} = \mathbf{b} \end{array},$$

where $\mathbf{P} \in \mathbb{S}_+^n$, $\mathbf{q} \in \mathbb{R}^n$, $r \in \mathbb{R}$, $\mathbf{G} \in \mathbb{R}^{m\times n}$, $\mathbf{h} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{p\times n}$, and $\mathbf{b} \in \mathbb{R}^p$. Here, we minimize a convex quadratic function over a polyhedron.



If the objective and constraints are convex quadratic functions, we get a *quadratically constrained quadratic program* (QCQP). This is given by

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\ \text{subject to} & (1/2)\,\mathbf{x}^T\mathbf{P}_k\mathbf{x} + \mathbf{q}_k^T\mathbf{x} + r_k \leq 0\,,\ k = 1,\ldots,m\,,\ \mathbf{A}\mathbf{x} = \mathbf{b} \end{array},$$

where $\mathbf{P}_k \in \mathbb{S}_+^n$, $\mathbf{q}_k \in \mathbb{R}^n$, and $r_k \in \mathbb{R}$ for $k = 0, 1, \ldots, m$, $\mathbf{A} \in \mathbb{R}^{p\times n}$, and $\mathbf{b} \in \mathbb{R}^p$. Here, we minimize a convex quadratic function over a polyhedron.

Examples of QPs

- **Least squares:**

$$\text{minimize} \quad ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2^2 \ .$$

  - This has an analytic solution given by $\mathbf{x}^{\star} = \mathbf{A}^{\#}\mathbf{b}$.
  - We can add linear constraints, such as $\mathbf{l} \preceq \mathbf{x} \preceq \mathbf{u}$ (box constraints).

- **Linear program with random cost:**

$$\begin{array}{ll} \text{minimize} & \overline{\mathbf{c}}^T\mathbf{x} + \gamma\mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x} = E\left[\mathbf{c}^T\mathbf{x}\right] + \gamma\mathrm{Var}\left(\mathbf{c}^T\mathbf{x}\right) \\ \text{subject to} & \mathbf{G}\mathbf{x} \preceq \mathbf{h}\,,\ \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

  - Here, $\mathbf{c}$ is a random vector with mean $\overline{\mathbf{c}}$ and covariance $\boldsymbol{\Sigma}$.
  - Thus, $\mathbf{c}^T\mathbf{x}$ is a random variable with mean $\overline{\mathbf{c}}^T\mathbf{x}$ and variance $\mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x}$.
  - The quantity $\gamma > 0$ is a *risk aversion parameter* in that it controls the trade-off between the expected cost and the variance or risk.

Examples of QPs (Continued)

- **Distance between polyhedra:**
  The Euclidean distance between the polyhedra $\mathcal{P}_1 = \{\mathbf{x} : \mathbf{A}_1\mathbf{x} \preceq \mathbf{b}_1\}$ and $\mathcal{P}_2 = \{\mathbf{x} : \mathbf{A}_2\mathbf{x} \preceq \mathbf{b}_2\}$ is given by

$$\mathrm{dist}(\mathcal{P}_1, \mathcal{P}_2) = \inf \left\{ ||\mathbf{x}_1 - \mathbf{x}_2||_2 : \mathbf{x}_1 \in \mathcal{P}_1 \, , \; \mathbf{x}_2 \in \mathcal{P}_2 \right\} \, ,$$

  and can be found by solving the QP

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x}_1 - \mathbf{x}_2||_2^2 \\ \text{subject to} & \mathbf{A}_1\mathbf{x}_1 \preceq \mathbf{b}_1 \, , \; \mathbf{A}_2\mathbf{x}_2 = \mathbf{b}_2 \end{array} .$$

- **Bounding variance:**
  As an extension to the Chebyshev inequalities considered above, suppose we want to bound the *variance* σ^2 of a function f of a discrete random variable x:

$$\sigma^2 = E\left[f^2(x)\right] - \left(E[f(x)]\right)^2 = \sum_{k=1}^{n} f_k^2 p_k - \left(\sum_{k=1}^{n} f_k p_k \right)^2 \, ,$$

  with $f_k \triangleq f(u_k)$. This is a *concave* quadratic function of $\mathbf{p}$. So, we can *maximize* the variance of $f(x)$, subject to the given prior information, by solving the QP

$$\begin{array}{ll} \text{minimize} & \sum_{k=1}^{n} f_k^2 p_k - \left(\sum_{k=1}^{n} f_k p_k\right)^2 \\ \text{subject to} & \mathbf{p} \succeq \mathbf{0} \, , \; \mathbf{1}^T\mathbf{p} = 1 \, , \; \alpha_k \leq \mathbf{a}_k^T\mathbf{p} \leq \beta_k \, , \; k = 1, \ldots, m \end{array} .$$

Second-Order Cone Programs

A problem closely related to quadratic programming is the *second-order cone program* (SOCP) given by

$$\begin{array}{ll} \text{minimize} & \mathbf{f}^T\mathbf{x} \\ \text{subject to} & ||\mathbf{A}_k\mathbf{x}+\mathbf{b}_k||_2 \leq \mathbf{c}_k^T\mathbf{x}+d_k\,,\ k=1,\ldots,m \\ & \mathbf{F}\mathbf{x}=\mathbf{g} \end{array},$$

where $\mathbf{f} \in \mathbb{R}^n$, for $k = 1, \ldots, m$ we have $\mathbf{A}_k \in \mathbb{R}^{n_k \times n}$, $\mathbf{b}_k \in \mathbb{R}^{n_k}$, $\mathbf{c}_k \in \mathbb{R}^n$, and $d_k \in \mathbb{R}$, and finally $\mathbf{F} \in \mathbb{R}^{p \times n}$ and $\mathbf{g} \in \mathbb{R}^p$.

- The inequalities are called *second-order cone (SOC) constraints*, since we have

$$\left(\mathbf{A}_k\mathbf{x}+\mathbf{b}_k, \mathbf{c}_k^T\mathbf{x}+d_k\right) \in \text{second-order cone in } \mathbb{R}^{n_k+1}\,.$$

- If $n_k = 0$ for $k = 1\ \ldots, m$, it reduces to an LP. Similarly, if $\mathbf{c}_k = \mathbf{0}$ for $k = 1, \ldots, m$, it reduces to a QCQP.
- It is more general than LPs and QCQPs.

Robust Linear Programming: Deterministic Case

Parameters in optimization problems often come with some uncertainty. For example, in an LP

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x} \leq b_k \ , \ k = 1, \ldots, m \end{array} ,$$

there may be uncertainty in $\mathbf{c}$, $\mathbf{a}_k$, or b_k. For example, suppose $\mathbf{c}$ and b_k are fixed, and each vector $\mathbf{a}_k$ is uncertain, but can be deterministically bounded to lie with within a given ellipsoid $\mathcal{E}_k$ (called the *uncertainty ellipsoid*). Namely, we have

$$\mathbf{a}_k \in \mathcal{E}_k = \left\{ \overline{\mathbf{a}}_k + \mathbf{P}_k \mathbf{u} : ||\mathbf{u}||_2 \leq 1 \right\} ,$$

where $\overline{\mathbf{a}}_k \in \mathbb{R}^n$ is the *center* of the ellipsoid and the singular values/vectors of $\mathbf{P} \in \mathbb{R}^{n \times n}$ dictate the semi-axis length/orientations of the ellipsoid, respectively. This leads to the *robust LP*:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{a}_k^T \mathbf{x} \leq b_k \ , \ \mathbf{a}_k \in \mathcal{E}_k \ , \ k = 1, \ldots, m \end{array} .$$

Since $\sup_{||\mathbf{u}||_2 \leq 1} \left\{ \left(\overline{\mathbf{a}}_k + \mathbf{P}_k \mathbf{u} \right)^T \mathbf{x} \right\} = \overline{\mathbf{a}}_k^T \mathbf{x} + \left|\left| \mathbf{P}_k^T \mathbf{x} \right|\right|_2$, the robust LP above is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \overline{\mathbf{a}}_k^T \mathbf{x} + \left|\left| \mathbf{P}_k^T \mathbf{x} \right|\right|_2 \leq b_k \ , \ \mathbf{a}_k \in \mathcal{E}_k \ , \ k = 1, \ldots, m \end{array} .$$

Robust Linear Programming: Stochastic Case

- Here, we assume that $\mathbf{a}_k$ is now a random vector. Specifically, we assume $\mathbf{a}_k$ is *Gaussian* with mean $\overline{\mathbf{a}}_k$ and covariance $\mathbf{\Sigma}_k$ (i.e., $\mathbf{a}_k \sim \mathcal{N}(\overline{\mathbf{a}}_k, \mathbf{\Sigma}_k)$).
- As such, $\mathbf{a}_k^T \mathbf{x}$ is a Gaussian random variable with mean $\overline{\mathbf{a}}_k^T \mathbf{x}$ and variance $\mathbf{x}^T \mathbf{\Sigma}_k \mathbf{x}$. Hence, we have

$$\Pr\left\{\mathbf{a}_k^T \mathbf{x} \leq b_k\right\} = \Phi\left(\frac{b_k - \overline{\mathbf{a}}_k^T \mathbf{x}}{\left\|\mathbf{\Sigma}_k^{1/2}\mathbf{x}\right\|_2}\right) ,$$

where $\Phi(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{x} e^{-t^2/2}\, dt$ is the cumulative distribution function (cdf) of the standard normal distribution $\mathcal{N}(0,1)$.
- This leads to the robust LP

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \Pr\left\{\mathbf{a}_k^T \mathbf{x} \leq b_k\right\} \geq \eta \,,\ k = 1, \ldots, m \end{array} ,$$

where η is a *confidence* parameter.
- When $\eta \geq 1/2$, this is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \overline{a}_k^T \mathbf{x} + \Phi^{-1}(\eta) \left\|\mathbf{\Sigma}_k^{1/2}\mathbf{x}\right\|_2 \leq b_k \,,\ k = 1, \ldots, m \end{array} .$$