Convex optimization problems (II)

Lijun Zhang
zlj@nju.edu.cn
http://cs.nju.edu.cn/zlj

Outline

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

Linear Optimization Problems

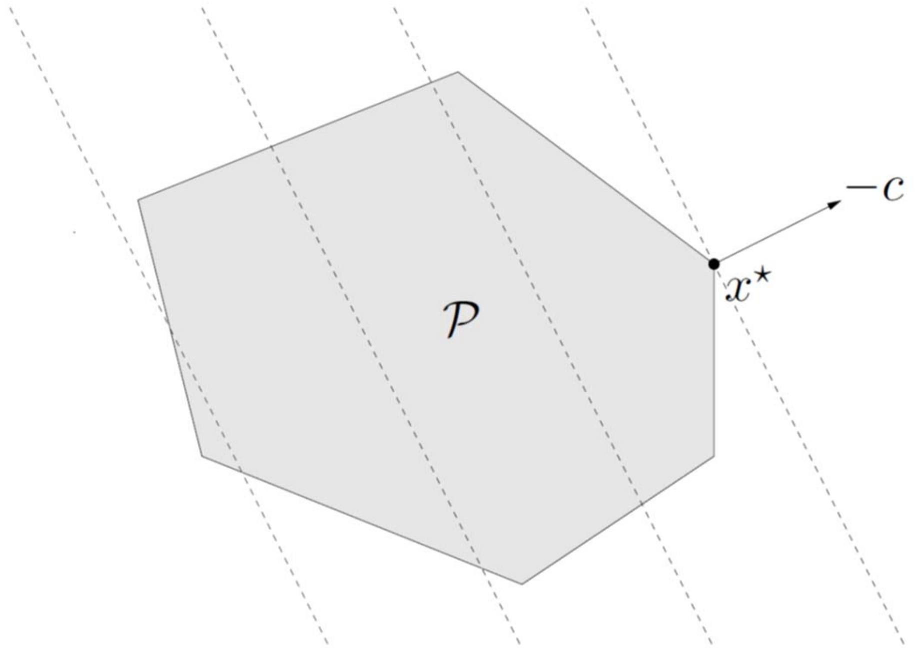
Linear Program (LP)

$$\begin{aligned} \min \quad & c^\top x + d \\ \text{s.t.} \quad & Gx \preccurlyeq h \\ & Ax = b \end{aligned}$$

- $G \in \mathbf{R}^{m \times n}$ and $A \in \mathbf{R}^{p \times n}$
- It is common to omit the constant d
- Maximization problem with affine objective and constraint functions is also an LP
- The feasible set of LP is a polyhedron $\mathcal{P}$

Linear Optimization Problems

□ Geometric Interpretation of an LP



- The objective $c^\top x$ is linear, so its level curves are hyperplanes orthogonal to c
- x^* is as far as possible in the direction $-c$

Two Special Cases of LP

- Standard Form LP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \\ & x \succcurlyeq 0 \end{aligned}$$

  - The only inequalities are $x \succcurlyeq 0$

- Inequality Form LP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax \preccurlyeq b \end{aligned}$$

  - No equality constraint

Converting to Standard Form

Conversion

$$\begin{array}{ll} \min & c^\top x + d \\ \text{s.t.} & Gx \preccurlyeq h \\ & Ax = b \end{array} \quad \Rightarrow \quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Ax = b \\ & x \succcurlyeq 0 \end{array}$$

- To use an algorithm for standard LP

Introduce Slack Variables s

$$\begin{array}{ll} \min & c^\top x + d \\ \text{s.t.} & Gx \preccurlyeq h \\ & Ax = b \end{array} \quad \Rightarrow \quad \begin{array}{ll} \min & c^\top x + d \\ \text{s.t.} & Gx + s = h \\ & Ax = b \\ & s \succcurlyeq 0 \end{array}$$

Converting to Standard Form

- Decompose x

$$x = x^+ - x^-, \qquad x^+, x^- \geqslant 0$$

- Standard Form LP

$$\begin{aligned} \min \quad & c^\top x + d \\ \text{s.t.} \quad & Gx + s = h \\ & Ax = b \\ & s \geqslant 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \min \quad & c^\top x^+ - c^\top x^- + d \\ \text{s.t.} \quad & Gx^+ - Gx^- + s = h \\ & Ax^+ - Ax^- = b \\ & x^+ \geqslant 0, x^- \geqslant 0, s \geqslant 0 \end{aligned}$$

Example

Diet Problem

- Choose nonnegative quantities $x_1, \ldots, x_n$ of n foods
- One unit of food j contains amount a_{ij} of nutrient i, and costs c_j
- Healthy diet requires nutrient i in quantities at least b_i
- Determine the cheapest diet that satisfies the nutritional requirements

$$
\begin{aligned}
\min \quad & c^\top x \\
\text{s.t.} \quad & Ax \geqslant b \\
& x \geqslant 0
\end{aligned}
$$

Example

Chebyshev Center of a Polyhedron

- Find the largest Euclidean ball that lies in the polyhedron

$$\mathcal{P} = \{x \in \mathbf{R}^n | a_i^\top x \leq b_i, i = 1, \dots, m\}$$

- The center of the optimal ball is called the Chebyshev center of the polyhedron
- Represent the ball as $\mathcal{B} = \{x_c + u | \|u\|_2 \leq r\}$
- $x_c \in \mathbf{R}^n$ and r are variables, and we wish to maximize r subject to $\mathcal{B} \in \mathcal{P}$
- $\forall x \in \mathcal{B}, a_i^\top x \leq b_i \Leftrightarrow a_i^\top (x_c + u) \leq b_i, \|u\|_2 \leq r \Leftrightarrow a_i^\top x_c + \sup\{a_i^\top u | \|u\|_2 \leq r\} \leq b_i \Leftrightarrow a_i^\top x_c + r\|a_i\|_2 \leq b_i$

Example

Chebyshev Center of a Polyhedron

- Find the largest Euclidean ball that lies in the polyhedron

$$\mathcal{P} = \{x \in \mathbf{R}^n \mid a_i^\top x \le b_i, i = 1, \dots, m\}$$

- The center of the optimal ball is called the Chebyshev center of the polyhedron
- Represent the ball as $\mathcal{B} = \{x_c + u \mid \|u\|_2 \le r\}$
- $x_c \in \mathbf{R}^n$ and r are variables, and we wish to maximize r subject to $\mathcal{B} \in \mathcal{P}$

$$\begin{aligned} \max \quad & r \\ \text{s.t.} \quad & a_i^\top x_c + r\|a_i\|_2 \le b_i, \qquad i = 1, \dots, m \end{aligned}$$

Example

Chebyshev Inequalities

- x is a random variable on $\{u_1, \dots, u_n\} \subseteq \mathbf{R}$
- $p_i = \mathbf{prob}(x = u_i)$, $p \succcurlyeq 0, \mathbf{1}^\top p = 1$
- $\mathbf{E}f(x) = \sum_{i=1}^n p_i f(u_i)$ is a linear function of p
- Prior knowledge is given as

$$\alpha_i \le a_i^\top p \le \beta_i, \qquad i = 1, \dots, m$$

- To find a lower bound of $\mathbf{E}f_0(x) = a_0^\top p$

$$\begin{aligned} \min \quad & a_0^\top p \\ \text{s.t.} \quad & p \succcurlyeq 0, \mathbf{1}^\top p = 1 \\ & \alpha_i \le a_i^\top p \le \beta_i, \qquad i = 1, \dots, m \end{aligned}$$

Example

□ Piecewise-linear Minimization

- Consider the (unconstrained) problem

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

- The epigraph problem

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \max_{i=1,\dots,m} (a_i^\top x + b_i) \le t \end{aligned}$$

- An LP problem

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & a_i^\top x + b_i \le t, \qquad i = 1, \dots, m \end{aligned}$$

Linear-fractional Programming

Linear-fractional Program

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & Gx \preccurlyeq h \\ & Ax = b \end{aligned}$$

- The objective function is a ratio of affine functions

$$f_0(x) = \frac{c^\top x + d}{e^\top x + f}$$

- The domain is

$$\operatorname{dom} f_0 = \{x | e^\top x + f > 0\}$$

- A quasiconvex optimization problem

Linear-fractional Programming

Transforming to a linear program

$$\begin{aligned} \min \quad & f_0(x) = \frac{c^\top x + d}{e^\top x + f} \\ \text{s.t.} \quad & Gx \preccurlyeq h \\ & Ax = b \end{aligned} \quad \Rightarrow \quad \begin{aligned} \min \quad & c^\top y + dz \\ \text{s.t.} \quad & Gy - hz \preccurlyeq 0 \\ & Ay - bz = 0 \\ & e^\top y + fz = 1 \\ & z \geq 0 \end{aligned}$$

- **Proof**

x is feasible in LFP $\Rightarrow y = \frac{x}{e^\top x + f},\ z = \frac{1}{e^\top x + f}$ is feasible in LP, $c^\top y + dz = f_0(x) \Rightarrow$ the optimal value of LFP is greater than or equal to the optimal value of LP

Linear-fractional Programming

Transforming to a linear program

$$\begin{array}{ll} \min & f_0(x) = \dfrac{c^\top x + d}{e^\top x + f} \\ \text{s.t.} & Gx \preccurlyeq h \\ & Ax = b \end{array} \quad \Rightarrow \quad \begin{array}{ll} \min & c^\top y + dz \\ \text{s.t.} & Gy - hz \preccurlyeq 0 \\ & Ay - bz = 0 \\ & e^\top y + fz = 1 \\ & z \geq 0 \end{array}$$

- Proof

(y, z) is feasible in LP and $z \neq 0 \Rightarrow x = y/z$ is feasible in LFP, $f_0(x) = c^\top y + dz \Rightarrow$ the optimal value of LFP is less than or equal to the optimal value of LP

(y, z) is feasible in LP, $z = 0$ and x_0 is feasible in LFP $\Rightarrow x = x_0 + ty$ is feasible in LFP for all $t \geq 0$, $\lim_{t \to \infty} f_0(x_0 + ty) = c^\top y + dz$

Generalized Linear-fractional Programming

- Generalized Linear-fractional Program

$$f_0(x) = \max_{i=1,\dots,r} \frac{c_i^\top x + d_i}{e_i^\top x + f_i}$$

  - $\mathrm{dom}\, f_0 = \{x | e_i^\top x + f_i > 0, i = 1, \dots, r\}$
  - A quasiconvex optimization problem

- Von Neumann Growth Problem

$$\begin{array}{ll} \max & \min_{i=1,\dots,n} x_i^+ / x_i \\ \text{s.t.} & x^+ \succcurlyeq 0 \\ & Bx^+ \preccurlyeq Ax \end{array}$$

Generalized Linear-fractional Programming

Von Neumann Growth Problem

$$\begin{array}{ll} \max & \min\limits_{i=1,\ldots,n} x_i^+/x_i \\ \text{s.t.} & x^+ \succcurlyeq 0 \\ & Bx^+ \preccurlyeq Ax \end{array}$$

- $x, x^+ \in \mathbf{R}^n$: activity levels of n sectors, in current and next period
- $(Ax)_i, (Bx^+)_i$: produced and consumed amounts of good i
- $Bx^+ \preccurlyeq Ax$: goods consumed in the next period cannot exceed the goods produced in the current period
- x_i^+/x_i growth rate of sector i

Outline

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

Quadratic Optimization Problems

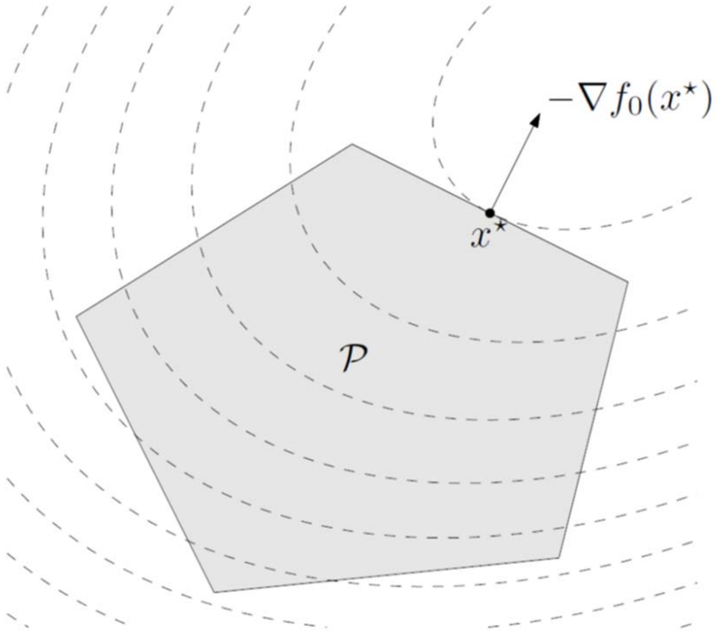
Quadratic Program (QP)

$$\begin{array}{ll} \min & (1/2)x^\top Px + q^\top x + r \\ \text{s.t.} & Gx \preccurlyeq h \\ & Ax = b \end{array}$$

- $P \in \mathbf{S}_+^n, G \in \mathbf{R}^{m\times n}$ and $A \in \mathbf{R}^{p\times n}$
- The objective function is (convex) quadratic
- The constraint functions are affine
- When $P = 0$, QP becomes LP

Quadratic Optimization Problems

Geometric Illustration of QP



- The feasible set $\mathcal{P}$ is a polyhedron
- The contour lines of the objective function are shown as dashed curves.

Quadratic Optimization Problems

- Quadratically Constrained Quadratic Program (QCQP)

$$
\begin{array}{lll}
\min & (1/2)x^\top P_0 x + q_0^\top x + r_0 & \\
\text{s.t.} & (1/2)x^\top P_i x + q_i^\top x + r_i \leq 0, & i = 1, \dots, m \\
& Ax = b &
\end{array}
$$

  - $P_i \in \mathbf{S}_+^n, i = 0, \dots, m$
  - The inequality constraint functions are (convex) quadratic
  - The feasible set is the intersection of ellipsoids (when $P_i \succ 0$) and an affine set
  - Include QP as a special case

Examples

- Least-squares and Regression

$$\min \ \|Ax-b\|_2^2 = x^\top A^\top A x - 2b^\top A x + b^\top b$$

  - Analytical solution: $x = A^\dagger b$
  - Can add linear constraints, e.g., $l \preccurlyeq x \preccurlyeq u$

- Distance Between Polyhedra

$$\begin{aligned} \min \quad & \|x_1 - x_2\|_2^2 \\ \text{s.t.} \quad & A_1 x_1 \preccurlyeq b_1, \qquad A_2 x_2 \preccurlyeq b_2 \end{aligned}$$

  - Find the distance between the polyhedra $\mathcal{P}_1 = \{x | A_1 x \preccurlyeq b_1\}$ and $\mathcal{P}_2 = \{x | A_2 x \preccurlyeq b_2\}$

$$\operatorname{dist}(\mathcal{P}_1, \mathcal{P}_2) = \inf\{\|x_1 - x_2\|_2 | x_1 \in \mathcal{P}_1, x_2 \in \mathcal{P}_2\}$$

Example

Bounding Variance

- x is a random variable on $\{u_1, \ldots, u_n\} \subseteq \mathbf{R}$
- $p_i = \mathbf{prob}(x = u_i)$, $p \succcurlyeq 0, \mathbf{1}^\top p = 1$
- The variance of a random variable $f(x)$

$$\mathbf{E}f^2 - (\mathbf{E}f)^2 = \sum_{i=1}^{n} f_i^2 p_i - \left(\sum_{i=1}^{n} f_i p_i\right)^2$$

- Maximize the variance

$$\begin{array}{ll} \max & \sum_{i=1}^{n} f_i^2 p_i - \left(\sum_{i=1}^{n} f_i p_i\right)^2 \\ \mathrm{s.t.} & p \succcurlyeq 0, \mathbf{1}^\top p = 1 \\ & \alpha_i \leq a_i^\top p \leq \beta_i, i = 1, \ldots, m \end{array}$$

Second-order Cone Programming

Second-order Cone Program (SOCP)

$$\begin{aligned} \min \quad & f^\top x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \le c_i^\top x + d_i, \qquad i = 1, \dots, m \\ & Fx = g \end{aligned}$$

- $A_i \in \mathbf{R}^{n_i \times n}, F \in \mathbf{R}^{p \times n}$
- Second-order Cone (SOC) constraint: $\|Ax + b\|_2 \le c^\top x + d$ where $A \in \mathbf{R}^{k \times n}$, is same as requiring $(Ax + b, c^\top x + d) \in \text{SOC in } \mathbf{R}^{k+1}$

$$\begin{aligned} \text{SOC} \quad & = \left\{ (x, t) \in \mathbf{R}^{k+1} \,|\, \|x\|_2 \le t \right\} \\ & = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \,\middle|\, \begin{bmatrix} x \\ t \end{bmatrix}^\top \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \le 0, t \ge 0 \right\} \end{aligned}$$

Second-order Cone Programming

- Second-order Cone Program (SOCP)

$$\begin{array}{lll} \min & f^\top x & \\ \text{s.t.} & \|A_i x + b_i\|_2 \le c_i^\top x + d_i, & i = 1, \dots, m \\ & Fx = g & \end{array}$$

  - $A_i \in \mathbf{R}^{n_i \times n}, F \in \mathbf{R}^{p \times n}$
  - Second-order Cone (SOC) constraint: $\|Ax + b\|_2 \le c^\top x + d$ where $A \in \mathbf{R}^{k \times n}$, is same as requiring $(Ax + b, c^\top x + d) \in \mathrm{SOC}$ in $\mathbf{R}^{k+1}$
  - If $c_i = 0, i = 1, \dots, m$, it reduces to QCQP by squaring each inequality constraint
  - More general than QCQP and LP

Example

Robust Linear Programming

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & a_i^\top x \le b_i, \qquad i = 1, \dots, m \end{aligned}$$

- There can be uncertainty in a_i
- Assume a_i are known to lie in ellipsoids

  $a_i \in \mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}$, $P_i \in R^{n\times n}$
- The constraints must hold for all $a_i \in \mathcal{E}_i$

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & a_i^\top x \le b_i \text{ for all } a_i \in \mathcal{E}_i, \qquad i = 1, \dots, m \end{aligned}$$

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \sup\{a_i^\top x \mid a_i \in \mathcal{E}_i\} \le b_i, \qquad i = 1, \dots, m \end{aligned}$$

Example

- Note that

$$\sup\{a_i^\top x \mid a_i \in \mathcal{E}_i\} = \bar{a}_i^\top x + \sup\{u^\top P_i^\top x \mid \|u\|_2 \leq 1\}$$
$$= \bar{a}_i^\top x + \left\|P_i^\top x\right\|_2$$

- Robust linear constraint

$$\bar{a}_i^\top x + \left\|P_i^\top x\right\|_2 \leq b_i$$

- SOCP

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & \bar{a}_i^\top x + \left\|P_i^\top x\right\|_2 \leq b_i, \qquad i = 1, \dots, m \end{array}$$

Example

- Linear Programming with Random Constraints
  - Suppose that a_i is independent Gaussian random vectors with mean $\bar{a}_i$ and covariance Σ_i
  - Require each constraint $a_i^\top x \le b_i$ holds with probability exceeding $\eta \ge 0.5$

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & \operatorname{prob}\left(a_i^\top x \le b_i\right) \ge \eta, \qquad i = 1, \dots, m \end{array}$$

Example

Linear Programming with Random Constraints (SOCP)

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \bar{a}_i^\top x + \Phi^{-1}(\eta) \left\| \Sigma_i^{1/2} x \right\|_2 \le b_i, \qquad i = 1, \dots, m \end{aligned}$$

- Analysis

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

$$\Phi\left(\frac{b_i - \bar{a}_i^\top x}{\left\| \Sigma_i^{1/2} x \right\|_2} \right) \ge \eta \Leftrightarrow \frac{b_i - \bar{a}_i^\top x}{\left\| \Sigma_i^{1/2} x \right\|_2} \ge \Phi^{-1}(\eta) \Leftrightarrow \bar{a}_i^T x +$$

$$\Phi^{-1}(\eta) \left\| \Sigma_i^{1/2} x \right\|_2 \le b_i$$

where $\Phi(z) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{z} e^{-t^2/2} dt$

Outline

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

Definitions

Monomial Function

$$f(x) = c x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$$

- $f: \mathbf{R}^n \to \mathbf{R}$, $\operatorname{dom} f = \mathbf{R}^n_{++}$, $c > 0$ and $a_i \in \mathbf{R}$
- Closed under multiplication, division, and nonnegative scaling

Posynomial Function

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \dots x_n^{a_{nk}}$$

- Closed under addition, multiplication, and nonnegative scaling

Geometric Programming (GP)

The Problem

$$
\begin{aligned}
\min \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \le 1, \qquad i = 1, \dots, m \\
& h_i(x) = 1, \qquad i = 1, \dots, p
\end{aligned}
$$

- $f_0, \dots, f_m$ are posynomials
- $h_1, \dots, h_p$ are monomials
- Domain of the problem

$$\mathcal{D} = \mathbf{R}^n_{++}$$

- Implicit constraint: $x \succ 0$

Extensions of GP

- f is a posynomial and h is a monomial

$$f(x) \le h(x) \Leftrightarrow \frac{f(x)}{h(x)} \le 1$$

- h_1 and h_2 are nonzero monomials

$$h_1(x) = h_2(x) \Leftrightarrow \frac{h_1(x)}{h_2(x)} = 1$$

- Maximize a nonzero monomial objective function by minimizing its inverse

$$\begin{array}{ll} \max & x/y \\ \text{s.t.} & 2 \le x \le 3 \\ & x^2 + 3y/z \le \sqrt{y} \\ & x/y = z^2 \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & x^{-1}y \\ \text{s.t.} & 2x^{-1} \le 1, (1/3)x \le 1 \\ & x^2 y^{-1/2} + y^{1/2} z^{-1} \le 1 \\ & x y^{-1} z^{-2} = 1 \end{array}$$

GP in Convex Form

- Change of Variables $y_i = \log x_i$
  - f is the monomial function

$$f(x) = c x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n}, \qquad x_i = e^{y_i}$$

$$f(x) = f(e^{y_1}, \ldots, e^{y_n}) = c(e^{y_1})^{a_1} \cdots (e^{y_n})^{a_n}$$

$$= e^{a_1 y_1 + \cdots + a_n y_n + \log c} = e^{a^\top y + b}$$

  - f is the posynomial function

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \ldots x_n^{a_{nk}}$$

$$f(x) = \sum_{k=1}^{K} e^{a_k^\top y + b_k}$$

GP in Convex Form

- New Form

$$\begin{aligned} \min \quad & \sum_{k=1}^{K_0} e^{a_{0k}^\top y + b_{0k}} \\ \text{s.t.} \quad & \sum_{k=1}^{K_i} e^{a_{ik}^\top y + b_{ik}} \leq 1, \qquad i = 1, \dots, m \\ & e^{g_i^\top y + h_i} = 1, \qquad i = 1, \dots, p \end{aligned}$$

- Taking the Logarithm

$$\begin{aligned} \min \quad & \tilde{f}_0(y) = \log\left(\sum_{k=1}^{K_0} e^{a_{0k}^\top y + b_{0k}}\right) \\ \text{s.t.} \quad & \tilde{f}_i(y) = \log\left(\sum_{k=1}^{K_i} e^{a_{ik}^\top y + b_{ik}}\right) \leq 0, \qquad i = 1, \dots, m \\ & \tilde{h}_i(y) = g_i^\top y + h_i = 0, \qquad i = 1, \dots, p \end{aligned}$$

Example

Frobenius Norm Diagonal Scaling

- Given a matrix $M \in \mathbf{R}^{n\times n}$
- Choose a diagonal matrix D such that DMD^{-1} is small

$$\|DMD^{-1}\|_F^2 = \operatorname{tr}\left((DMD^{-1})^\top (DMD^{-1})\right) = \sum_{i,j=1}^{n} (DMD^{-1})_{ij}^2$$

$$= \sum_{i,j=1}^{n} M_{ij}^2 d_i^2 / d_j^2$$

- Unconstrained GP

$$\min \quad \sum_{i,j=1}^{n} M_{ij}^2 d_i^2 / d_j^2$$

Outline

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

Generalized Inequality Constraints

- Convex Optimization Problem with Generalized Inequality Constraints

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \preccurlyeq_{K_i} 0, \qquad i = 1, \dots, m \\ & Ax = b \end{aligned}$$

  - $f_0: \mathbf{R}^n \to \mathbf{R}$ is convex;
  - $K_i \subseteq \mathbf{R}^{k_i}$ are proper cones
  - $f_i: \mathbf{R}^n \to \mathbf{R}^{k_i}$ is K_i-convex w.r.t. proper cone $K_i \subseteq \mathbf{R}^{k_i}$

Generalized Inequality Constraints

Convex Optimization Problem with Generalized Inequality Constraints

$$
\begin{aligned}
\min \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \preccurlyeq_{K_i} 0, \qquad i = 1, \dots, m \\
& Ax = b
\end{aligned}
$$

- The feasible set, any sublevel set, and the optimal set are convex
- Any locally optimal is globally optimal
- The optimality condition for differentiable f_0 holds without change

Conic Form Problems

Conic Form Problems

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Fx + g \preccurlyeq_K 0 \\ & Ax = b \end{aligned}$$

- A linear objective
- One inequality constraint function which is affine
- A generalization of linear programs

Conic Form Problems

- Conic Form Problems

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Fx + g \preccurlyeq_K 0 \\ & Ax = b \end{array}$$

- Standard Form

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & x \succcurlyeq_K 0 \\ & Ax = b \end{array}$$

- Inequality Form

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Fx + g \preccurlyeq_K 0 \end{array}$$

Semidefinite Programming

Semidefinite Program (SDP)

$$
\begin{aligned}
\min \quad & c^{\top} x \\
\text{s.t.} \quad & x_1 F_1 + \cdots + x_n F_n + G \preccurlyeq 0 \\
& Ax = b
\end{aligned}
$$

- $K = \mathbf{S}_+^k$
- $G, F_1, \ldots, F_n \in \mathbf{S}^k$ and $A \in \mathbf{R}^{p \times n}$
- Linear matrix inequality (LMI)
- If $G, F_1, \ldots, F_n$ are all diagonal, LMI is equivalent to a set of n linear inequalities, and SDP reduces to LP

Semidefinite Programming

Standard From SDP

$$
\begin{array}{ll}
\min & \mathrm{tr}(CX) \\
\mathrm{s.t.} & \mathrm{tr}(A_i X) = b_i, \qquad i = 1, \dots, p \\
& X \succcurlyeq 0
\end{array}
$$

- $X \in \mathbf{S}^n$ is the variable and $C, A_1, \dots, A_p \in \mathbf{S}^n$
- p linear equality constraints
- A nonnegativity constraint

Inequality Form SDP

$$
\begin{array}{ll}
\min & c^\top x \\
\mathrm{s.t.} & x_1 A_1 + \cdots + x_n A_n \preccurlyeq B
\end{array}
$$

- $B, A_1, \dots, A_p \in \mathbf{S}^k$ and no equality constraint

Semidefinite Programming

Multiple LMIs and Linear Inequalities

$$
\begin{array}{ll}
\min & c^\top x \\
\text{s.t.} & F^{(i)}(x) = x_1 F_1^{(i)} + \cdots + x_n F_n^{(i)} + G^{(i)} \preccurlyeq 0, i = 1, \dots, K \\
& Gx \preccurlyeq h, \qquad Ax = b
\end{array}
$$

- It is referred as SDP as well

Be transformed as

$$
\begin{array}{ll}
\min & c^\top x \\
\text{s.t.} & \operatorname{diag}\left(Gx - h, F^{(1)}(x), \dots, F^{(K)}(x)\right) \preccurlyeq 0 \\
& Ax = b
\end{array}
$$

- A standard SDP

Examples

Second-order Cone Programming

$$\begin{array}{lll} \min & c^\top x & \\ \text{s.t.} & \|A_i x + b_i\|_2 \le c_i^\top x + d_i, & i = 1, \dots, m \\ & Fx = g & \end{array}$$

- A conic form problem

$$\begin{array}{lll} \min & c^\top x & \\ \text{s.t.} & -(A_i x + b_i, c_i^\top x + d_i) \preccurlyeq_{K_i} 0, & i = 1, \dots, m \\ & Fx = g & \end{array}$$

in which

$$K_i = \{(y, t) \in \mathbf{R}^{n_i+1} \mid \|y\|_2 \le t\}$$

Example

Matrix Norm Minimization

$$\min \quad \|A(x)\|_2 = (\lambda_{\max}(A(x)^\top A(x)))^{1/2}$$

- $A(x) = A_0 + x_1A_1 + \cdots + x_nA_n$ and $A_i \in \mathbf{R}^{p\times q}$
- Fact: $\|A\|_2 \le t \Leftrightarrow A^\top A \preccurlyeq t^2 I$

A New Problem

$$\min \quad \|A(x)\|_2^2 \Leftrightarrow \begin{array}{ll} \min & s \\ \text{s.t.} & \|A(x)\|_2^2 \preccurlyeq s \end{array}$$

Example

Matrix Norm Minimization

$$\min \quad \|A(x)\|_2 = (\lambda_{\max}(A(x)^\top A(x)))^{1/2}$$

- $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ and $A_i \in \mathbf{R}^{p\times q}$
- Fact: $\|A\|_2 \leq t \Leftrightarrow A^\top A \preccurlyeq t^2 I$

A New Problem

$$\begin{array}{ll} \min & s \\ \text{s.t.} & A(x)^\top A(x) \preccurlyeq sI \end{array} \Leftrightarrow \begin{array}{ll} \min & s \\ \text{s.t.} & A(x)^\top A(x) - sI \preccurlyeq 0 \end{array}$$

- $A(x)^\top A(x) - sI$ is matrix convex

Example

Matrix Norm Minimization

$$\min \quad \|A(x)\|_2 = (\lambda_{\max}(A(x)^\top A(x)))^{1/2}$$

- $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ and $A_i \in \mathbf{R}^{p\times q}$
- Fact:

$$\|A\|_2 \le t \Leftrightarrow A^\top A \preccurlyeq t^2 I \Leftrightarrow \begin{bmatrix} tI & A \\ A^\top & tI \end{bmatrix} \succcurlyeq 0$$

SDP

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \begin{bmatrix} tI & A(x) \\ A(x)^\top & tI \end{bmatrix} \succcurlyeq 0 \end{aligned}$$

- A single linear matrix inequality

Outline

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

General and Convex Vector Optimization Problems

General Vector Optimization Problem

$$\begin{array}{lll} \min\ (\text{w.r.t. } K) & f_0(x) & \\ \text{s.t.} & f_i(x) \le 0, & i = 1, \dots, m \\ & h_i(x) = 0, & i = 1, \dots, p \end{array}$$

- $f_0: \mathbf{R}^n \to \mathbf{R}^q$ is a vector-valued objective function
- $K \in \mathbf{R}^q$ is a proper cone, which is used to compare objective values
- $f_i: \mathbf{R}^n \to \mathbf{R}$ are the inequality constraint functions
- $h_i: \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

General and Convex Vector Optimization Problems

- Convex Vector Optimization Problem

$$\begin{array}{lll} \min \text{ (w.r.t. } K) & f_0(x) & \\ \text{s.t.} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

  - $f_0: \mathbf{R}^n \to \mathbf{R}^q$ is K-convex
  - $f_i: \mathbf{R}^n \to \mathbf{R}$ are convex
  - $h_i: \mathbf{R}^n \to \mathbf{R}$ are affine

- x is better than or equal to y

$$f_0(x) \preccurlyeq_K f_0(y)$$

  - Could be incomparable

Optimal Points and Values

- Achievable Objective Values

$$\mathcal{O} = \{f_0(x) | \exists x \in \mathcal{D}, f_i(x) \le 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$$

- If $\mathcal{O}$ has a minimum element $f_0(x)$
  - x is optimal and $f_0(x)$ is the optimal value
- x^* is optimal if and only if it is feasible and

$$\mathcal{O} \subseteq f_0(x^\star) + K$$

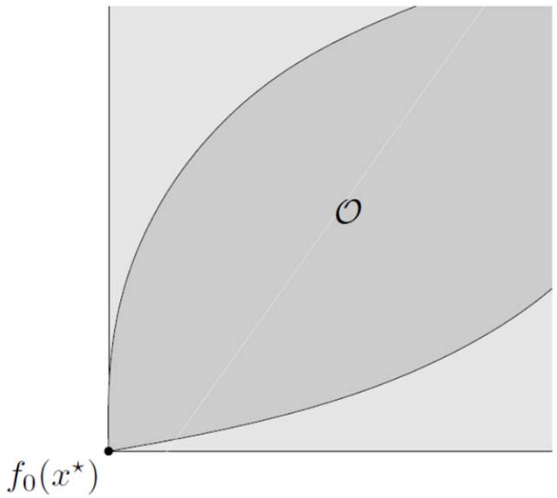
Optimal Points and Values

- Achievable Objective Values

$$\mathcal{O} = \{f_0(x) \mid \exists x \in \mathcal{D}, f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p\}$$

- If $\mathcal{O}$ has a minimum element $f_0(x)$
  - x is optimal and $f_0(x)$ is the optimal value
- $K = \mathbf{R}_+^2$

$$\mathcal{O} \subseteq f_0(x^\star) + K$$



Example

Best Linear Unbiased Estimator

- Suppose that $y = Ax + v$, where $v \in \mathbf{R}^m$ is noise, $y \in \mathbf{R}^m$ and $x \in \mathbf{R}^n$
- Estimate x from A and y
- Assume that A has rank n, and $\mathbf{E}v = 0, \mathbf{E}vv^\top = I$
- A linear estimator $\hat{x} = Fy$
- If $FA = I$, $\hat{x} = Fy$ is an unbiased linear estimator of x

$$\mathbf{E}\hat{x} = \mathbf{E}[FAx + Fv] = x$$

Example

- Best Linear Unbiased Estimator
  - The error covariance of an unbiased estimator

$$\mathbf{E}(\hat{x}-x)(\hat{x}-x)^\top = \mathbf{E}Fvv^\top F^\top = FF^\top$$

- Minimize the covariance

$$\begin{array}{ll} \min\ (\text{w.r.t. } \mathbf{S}_+^n) & FF^\top \\ \text{s.t.} & FA = I \end{array}$$

  - Solution

$$F^\star = A^\dagger = (A^\top A)^{-1}A^\top$$

$$F^\star {F^\star}^\top = (A^\top A)^{-1}$$

Pareto Optimal Points and Values

- Achievable Objective Values

$$\mathcal{O} = \{f_0(x) | \exists x \in \mathcal{D}, f_i(x) \le 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$$

- $f_0(x)$ is a minimal element of $\mathcal{O}$
  - x is Pareto optimal
  - $f_0(x)$ is a Pareto optimal value
- x is Pareto optimal if and only if it is feasible and

$$(f_0(x) - K) \cap \mathcal{O} = \{f_0(x)\}$$

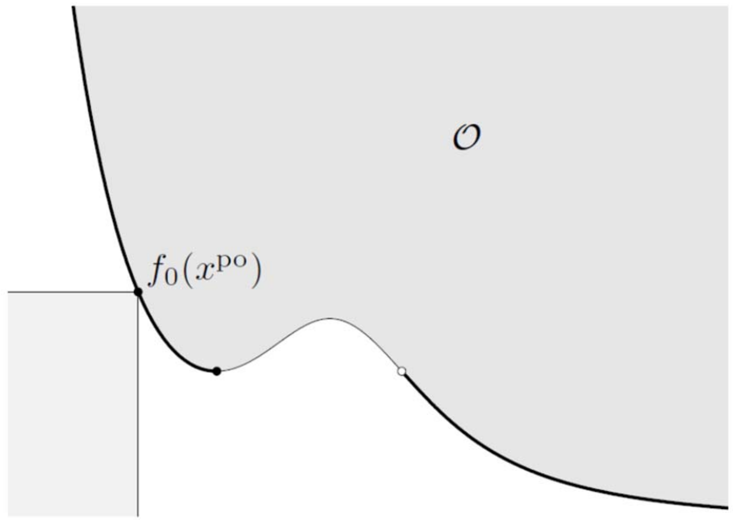
Pareto Optimal Points and Values

- Achievable Objective Values

$$\mathcal{O} = \{f_0(x) | \exists x \in \mathcal{D}, f_i(x) \le 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$$

- $f_0(x)$ is a minimal element of $\mathcal{O}$
  - x is Pareto optimal
  - $f_0(x)$ is a Pareto optimal value
- $K = \mathbf{R}_+^2$

$$(f_0(x) - K) \cap \mathcal{O} = \{f_0(x)\}$$



Pareto Optimal Points and Values

- Achievable Objective Values

$$\mathcal{O} = \{f_0(x) | \exists x \in \mathcal{D}, f_i(x) \le 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$$

- $f_0(x)$ is a minimal element of $\mathcal{O}$
  - x is Pareto optimal
  - $f_0(x)$ is a Pareto optimal value
- x is Pareto optimal if and only if it is feasible and

$$(f_0(x) - K) \cap \mathcal{O} = \{f_0(x)\}$$

- Let $\mathcal{P}$ be the set of Pareto optimal values

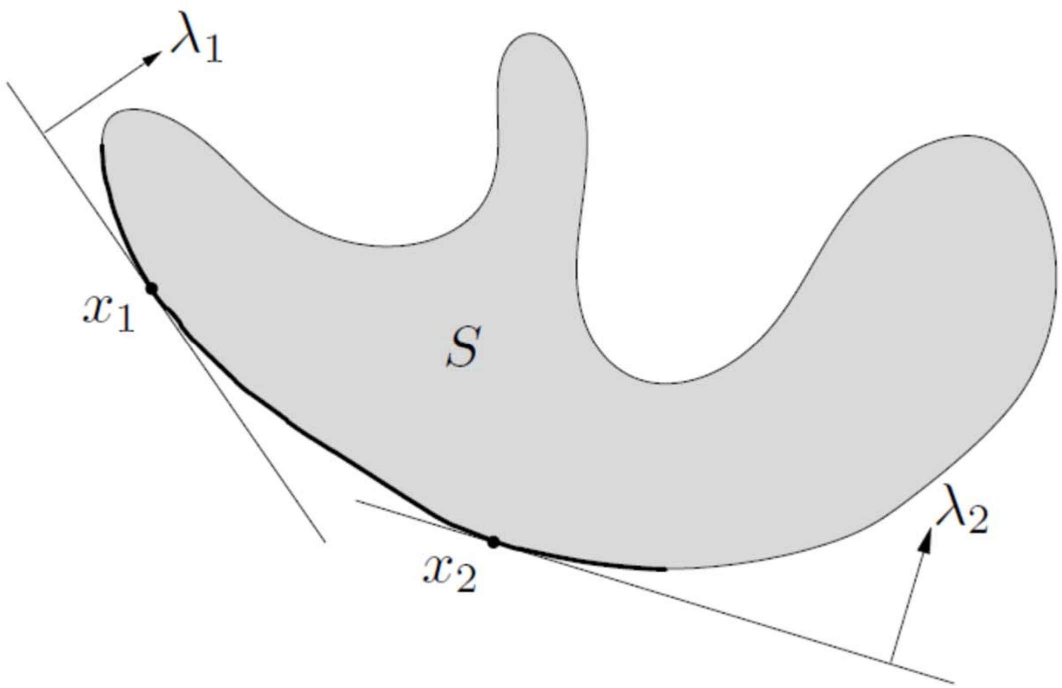
$$P \subseteq \mathcal{O} \cap \mathrm{bd}\mathcal{O}$$

Scalarization

- A standard technique for finding Pareto optimal (or optimal) points
- Find Pareto optimal points for any vector optimization problem by solving the ordinary scalar optimization problem
- Characterization of minimum and minimal points via dual generalized inequalities

Dual Characterization of Minimal Elements (1)

- If $\lambda \succ_{K^*} 0$, and x minimizes $\lambda^\top z$ over $z \in S$, then x is minimal.



Scalarization

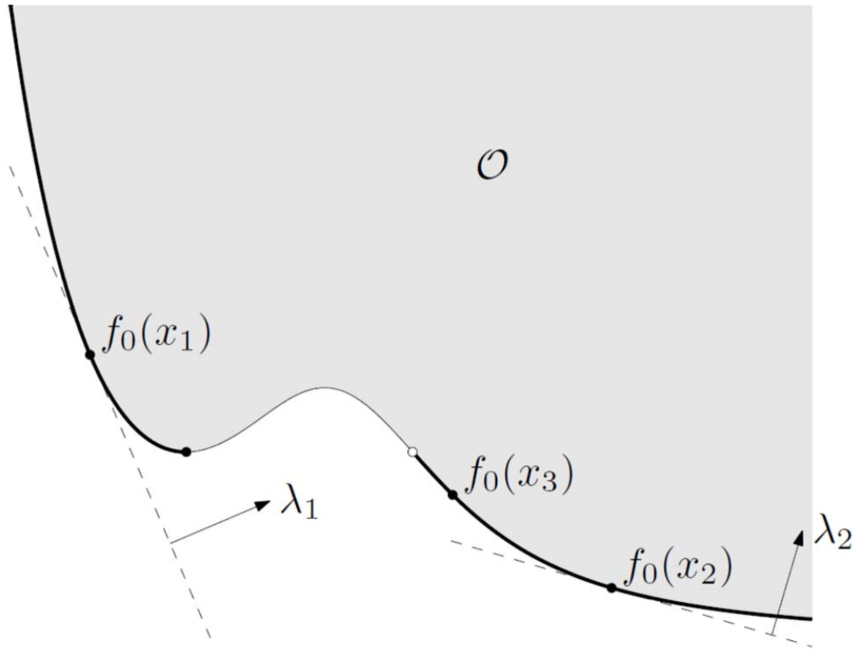
- Choose any $\lambda \succ_{K^*} 0$

$$\begin{aligned} \min \quad & \lambda^{\mathsf{T}} f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \qquad i = 1, \dots, m \\ & h_i(x) = 0, \qquad i = 1, \dots, p \end{aligned}$$

  - The optimal point x for this scalar problem is Pareto optimal for the vector optimization problem
  - λ is called the weight vector
  - By varying λ we obtain (possibly) different Pareto optimal solutions

Scalarization

- $K = \mathbf{R}_+^2$



- Scalarization cannot find $f_0(x_3)$

Scalarization of Convex Vector Optimization Problems

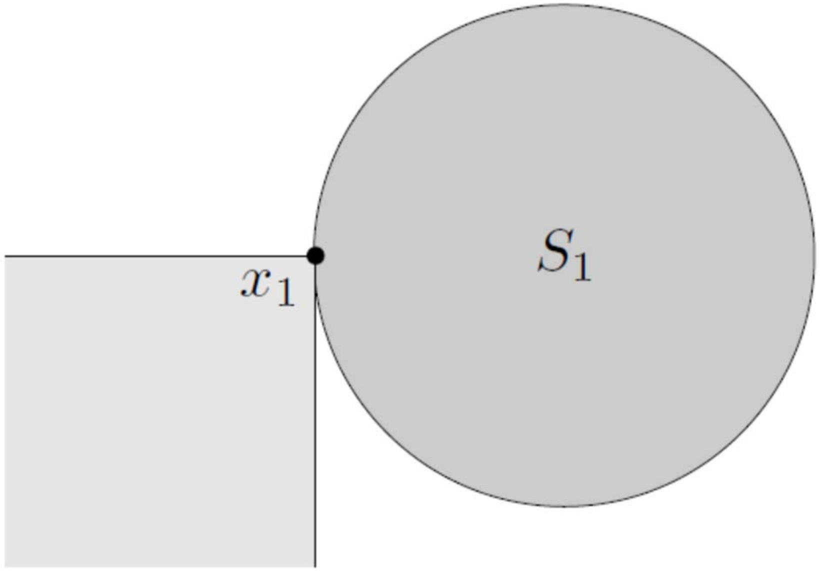
- Choose any $\lambda \succ_{K^*} 0$

$$\begin{array}{lll} \min & \lambda^{\mathsf{T}} f_0(x) & \\ \text{s.t.} & f_i(x) \le 0, & i = 1, \dots, m \\ & h_i(x) = 0, & i = 1, \dots, p \end{array}$$

  - A convex optimization problem
  - The optimal point x for this scalar problem is Pareto optimal for the vector optimization problem
  - λ is called the weight vector
  - By varying λ we obtain (possibly) different Pareto optimal solutions

Dual Characterization of Minimal Elements (2)

- If S is convex, for any minimal element x there exists a nonzero $\lambda \succcurlyeq_{K^*} 0$ such that x minimizes $\lambda^\top z$ over $z \in S$.



x_1 minimizes $\lambda^\top z$ over $z \in S_1$ for $\lambda = (1,0) \succcurlyeq 0$

Scalarization of Convex Vector Optimization Problems

- For every Pareto optimal point x^{po}, there is some nonzero $\lambda \succcurlyeq_{K^*} 0$ such that x^{po} is a solution of the scalarized problem

$$\begin{array}{lll} \min & \lambda^\top f_0(x) & \\ \text{s.t.} & f_i(x) \le 0, & i = 1, \dots, m \\ & h_i(x) = 0, & i = 1, \dots, p \end{array}$$

- It is not true that every solution of the scalarized problem, with $\lambda \succcurlyeq_{K^*} 0$ and $\lambda \neq 0$, is a Pareto optimal point for the vector problem

Scalarization of Convex Vector Optimization Problems

1. Consider all $\lambda \succ_{K^*} 0$

$$\begin{array}{lll} \min & \lambda^\top f_0(x) & \\ \text{s.t.} & f_i(x) \le 0, & i = 1, \dots, m \\ & h_i(x) = 0, & i = 1, \dots, p \end{array}$$

- Solve the above problem

2. Consider all $\lambda \succcurlyeq_{K^*} 0$, $\lambda \neq 0$, $\lambda \not\succ_{K^*} 0$
   - Solve the above problem
   - Verify the solution

Example

- Minimal Upper Bound on a Set of Matrices

$$\begin{array}{lll} \min\ (\text{w.r.t. } \mathbf{S}_+^n) & X & \\ \text{s.t.} & X \succcurlyeq A_i, & i = 1. \ldots, m \end{array}$$

  - $A_i \in \mathbf{S}^n, i = 1, \ldots, m$
  - The constraints mean that X is an upper bound on $A_1, \ldots, A_m$
  - A Pareto optimal solution is a minimal upper bound on the matrices

Example

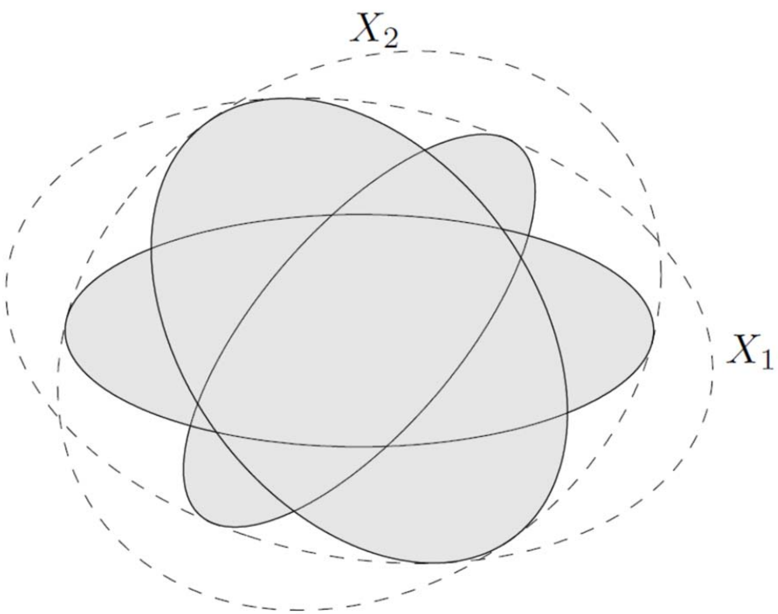
Scalarization

$$\begin{array}{lll} \min & \mathrm{tr}(WX) & \\ \mathrm{s.t.} & X \succcurlyeq A_i, & i = 1. \ldots, m \end{array}$$

- $W \in \mathbf{S}^n_{++}$
- An SDP
- If X is Pareto optimal for the vector problem then it is optimal for the SDP, for some nonzero weight matrix $W \succcurlyeq 0$.

Example

- A Simple Geometric Interpretation
  - Define an ellipsoid centered at the origin as $\mathcal{E}_A = \{u | u^\top A^{-1} u \leq 1\}$
  - $A \preccurlyeq B \Leftrightarrow \mathcal{E}_A \subseteq \mathcal{E}_B$



Multicriterion Optimization

- $K = \mathbf{R}_+^q$

$$f_0(x) = (F_1(x), \dots, F_q(x))$$

  - f_0 consists of q different objectives F_i and we want to minimize all F_i
  - It is convex if $f_1, \dots, f_m$ are convex, $h_1, \dots, h_p$ are affine, and $F_1, \dots, F_q$ are convex
  - Feasible $x^\star$ is optimal if

$$y \text{ is feasible} \Rightarrow f_0(x^\star) \preccurlyeq f_0(y)$$

  - Feasible x^{po} is Pareto optimal if

$$y \text{ is feasible},\ f_0(y) \preccurlyeq f_0(x^{\text{po}}) \Rightarrow f_0(x^{\text{po}}) = f_0(y)$$

Example

Regularized Least-Squares

$$\min\left(\text{w.r.t. } \mathbf{R}_+^2\right) \quad f_0(x) = (F_1(x), F_2(x))$$

- $F_1(x) = \|Ax - b\|_2^2$ measures the misfit
- $F_2(x) = \|x\|_2^2$ measures the size
- Our goal is to find x that gives a good fit and that is not large

Scalarization

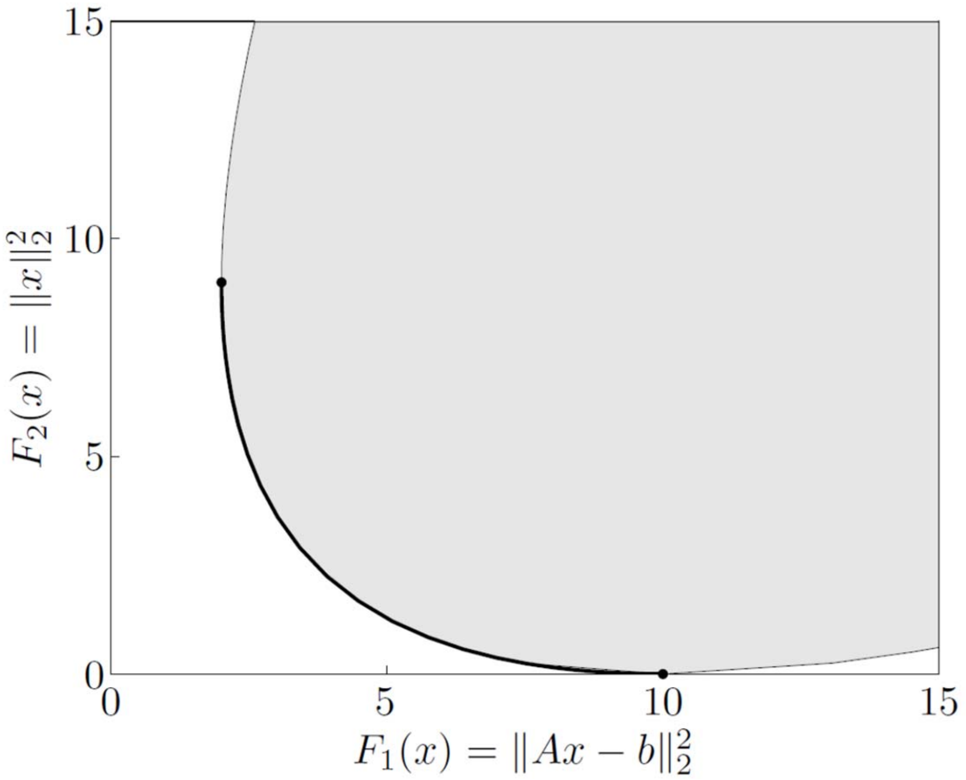
$$\begin{aligned}\lambda^\top f_0(x) &= \lambda_1 F_1(x) + \lambda_2 F_2(x) \\ &= x^\top(\lambda_1 A^\top A + \lambda_2 I)x - 2\lambda_1 b^\top Ax + \lambda_1 b^\top b\end{aligned}$$

Example

Solution

$$x(\mu) = (\lambda_1 A^\top A + \lambda_2 I)^{-1} \lambda_1 A^\top b = (A^\top A + \mu I)^{-1} A^\top b$$

- $\mu = \lambda_2 / \lambda_1$
- $\lambda = (0,1)$, we get $x = 0$
- With $\lambda \to (1,0)$, we get $x = A^\dagger b$



Summary

- Linear Optimization Problems
- Quadratic Optimization Problems
- Geometric Programming
- Generalized Inequality Constraints
- Vector Optimization

Homework 2

- http://www.lamda.nju.edu.cn/wanyy/optfall2020gra.html

- Deadline：Dec 8