COMP 9602: Convex Optimization

Convex Programs (III)

Dr. C Wu

Department of Computer Science
The University of Hong Kong

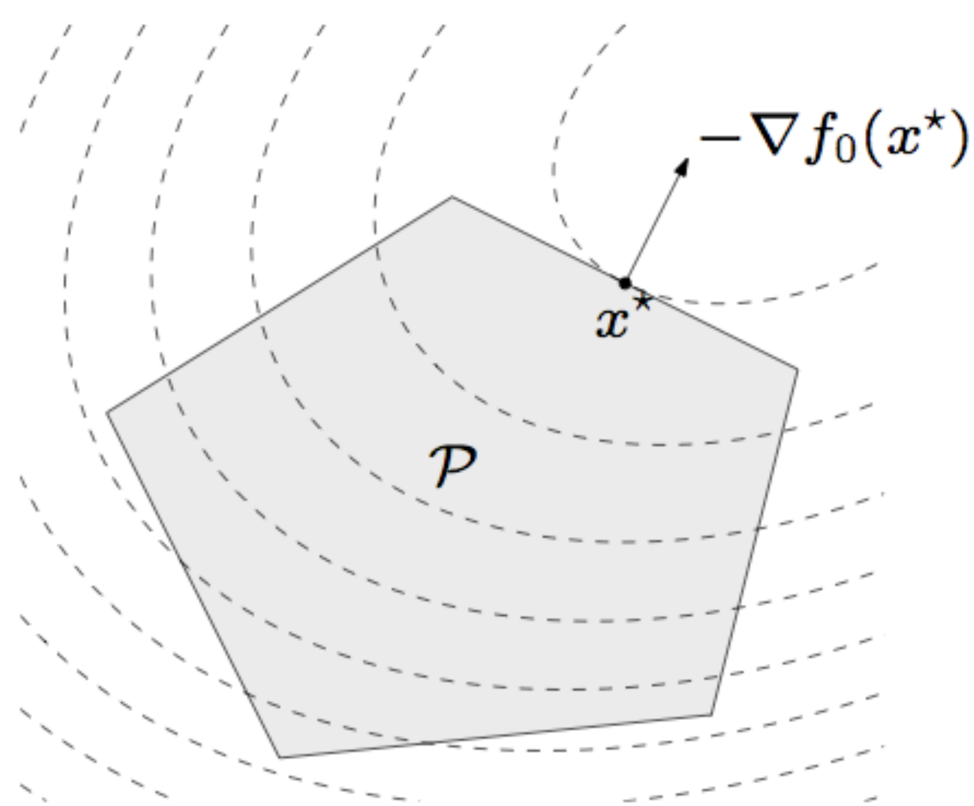
Where we are

Theory	convex set convex function standard forms of optimization problems, quasi-convex optimization linear program, integer linear program quadratic program geometric program semidefinite program vector optimization Duality
Algorithm	unconstrained optimization equality constrained optimization interior-point method localization methods subgradient method decomposition methods

Quadratic program (QP)

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

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



Examples

- Least squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

  - analytical solution $x^* = (A^T A)^{-1} A^T b$

- Linear program with random cost

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E}\, c^T x + \gamma\, \mathbf{var}(c^T x) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array}$$

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

Quadratically constrained quadratic programming (QCQP)

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- $P_i \in \mathbf{S}_+^n$; objective and constraints are convex quadratic
- if $P_1, \ldots, P_m \in \mathbf{S}_{++}^n$, feasible region is intersection of m ellipsoids and an affine set
- example

box constraints:

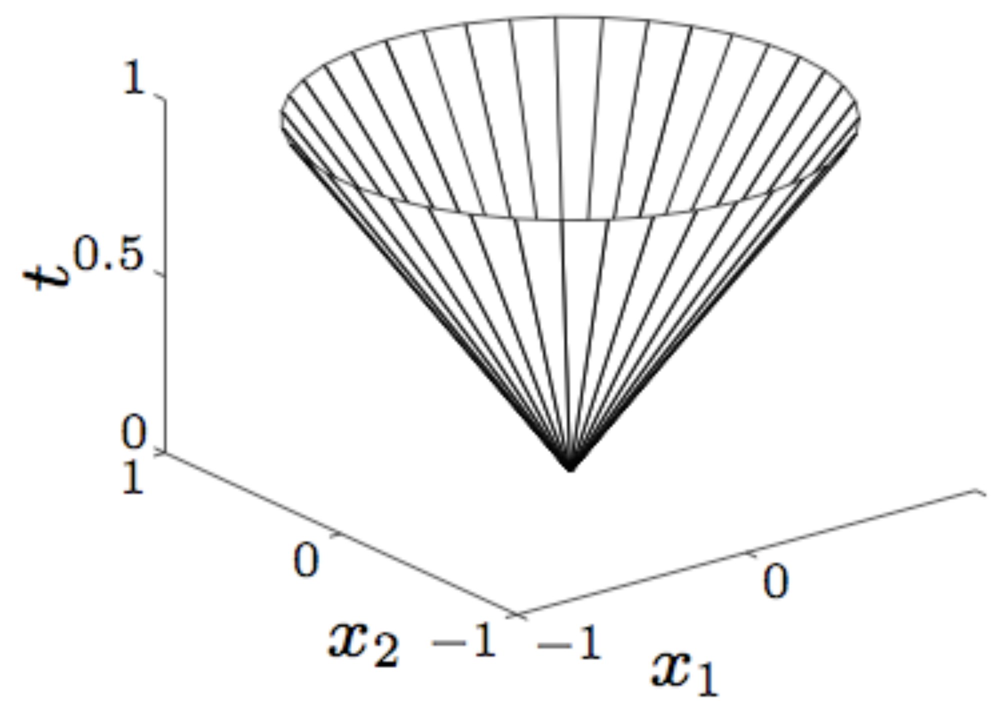
$$x \in [0,1]^n \Longleftrightarrow x_i(x_i - 1) \leq 0, i = 1, \ldots, n$$

Second-order cone programming (SOCP)

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

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

- Second-order cone
  - **norm cone:** $\{(x, t) \mid \|x\| \le t\}$
  - when the norm is Euclidean norm: second-order cone/ quadratic cone/ice-cream cone



Boundary of second-order cone in $\mathbf{R}^3$, $\{(x_1, x_2, t) | (x_1^2 + x_2^2)^{1/2} \le t\}$

- The inequality constraints are called second-order cone (SOC) constraints

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

Second-order cone programming (SOCP)

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

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

- More general than QCQP and LP
  - if $n_i=0$, reduces to LP
  - if $c_i=0$, reduces to QCQP

Robust linear programming

- The parameters in an LP can be uncertain
  - e.g.,

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

    there can be uncertainty in c, a_i, b_i

- Two common approaches to handle uncertainty in a_i (for simplicity)
  - convert to deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$

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

  - convert to stochastic model: constraints must hold with probability η (a_i as the random variable)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \ldots, m \end{array}$$

Robust linear programming

- Deterministic approach via SOCP
  - assume all a_i lie in given ellipsoids

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

center is $\bar{a}_i$, semi-axes determined by $\sqrt{\lambda_i}$, λ_i: eigenvalues of P_i^2

robust LP

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

is equivalent to the SOCP

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

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

Robust linear programming

Stochastic approach via SOCP

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

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

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

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \ldots, m, \end{array}$$

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

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

Geometric programming

monomial function

$$f(x) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

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

posynomial function: sum of monomials

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

geometric program (GP)

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

GP by itself is not convex

with f_i posynomial, h_i monomial

Geometric program in convex form

change variables to $y_i = \log x_i$, and take logarithm of objective, constraints

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

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

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

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

- geometric program transforms to convex problem

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

(Slide 19, Convex Function notes)

Example

- Maximize the volume of a box-shaped structure
  - optimization variables: height h, width w, depth d
  - constraints: a limit on the total wall area, a;
    a limit on the floor area, b;
    lower and upper bounds on the aspect ratio h/w, c and e;
    lower and upper bounds on the aspect ratio d/w, f and g

$$\max hwd$$

subject to:

$$2(hw + hd) \le a$$
$$wd \le b$$
$$c \le h/w \le e$$
$$f \le d/w \le g$$

not a GP

===>

$$\min h^{-1}w^{-1}d^{-1}$$

subject to:

$$(2/a)hw + (2/a)hd \le 1$$
$$(1/b)wd \le 1$$
$$ch^{-1}w \le 1$$
$$(1/e)hw^{-1} \le 1$$
$$fwd^{-1} \le 1$$
$$(1/g)w^{-1}d \le 1$$

GP

Semidefinite programming (SDP)

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\
& Ax = b
\end{array}
$$

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

- inequality constraint is called linear matrix inequality (LMI)

Example

- Minimize maximum eigenvalue

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

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

<=>

$$\text{minimize} \quad t$$

subject to:

$$\lambda_{max}(A(x)) \le t$$

<=> SDP

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

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

Vector optimization

general vector optimization problem

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

vector-valued objective function:

$f_0 : \mathbf{R}^n \to \mathbf{R}^q$, minimized w.r.t. proper cone $K \in \mathbf{R}^q$

i.e., find x^* such that $f_0(x^*) \preccurlyeq_K f_0(x)$, for all feasible x

$$f_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$$

$$h_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, p$$

Convex vector optimization

convex vector optimization problem

$$
\begin{array}{ll}
\text{minimize (w.r.t. } K) & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

with f_0 K-convex, f_1, . . . , f_m convex

K-convexity for vector-valued functions

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y)$$

for x, $y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$

Multicriterion optimization

- Or multi-objective optimization

vector optimization problem with $K = \mathbf{R}^q_+$

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

- q different objectives F_i; roughly speaking we want all F_i's to be small

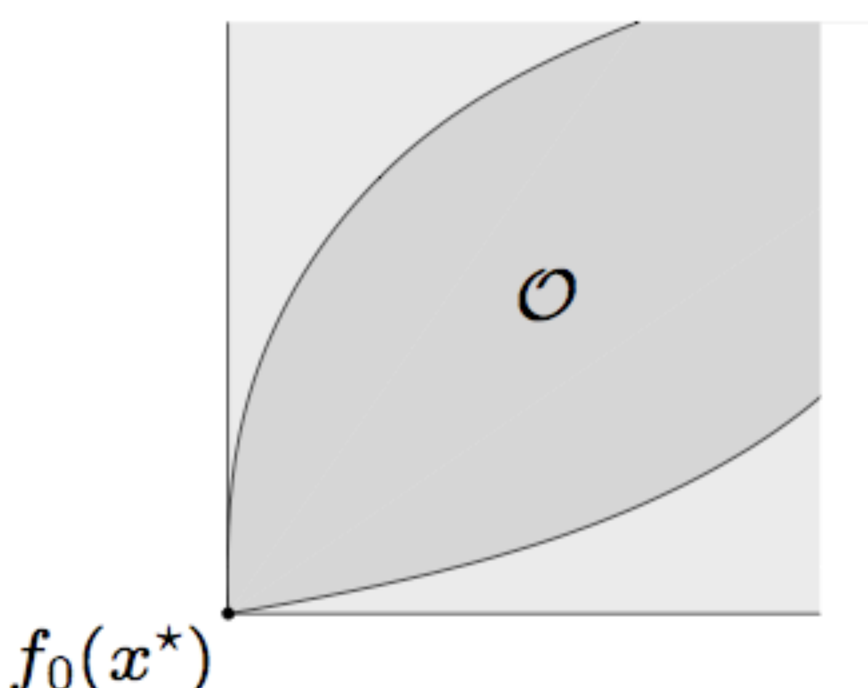
Optimal and Pareto optimal points

set of achievable objective values

$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- feasible x^* is **optimal** if $f_0(x^*)$ is a minimum value of $\mathcal{O}$

i.e., y feasible $\implies f_0(x^\star) \preceq_K f_0(y)$

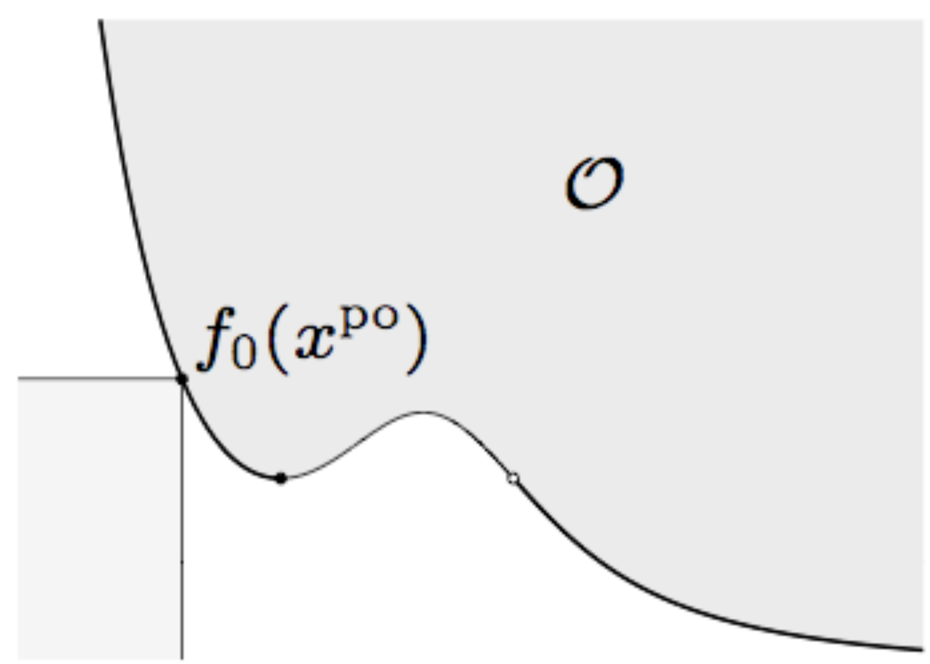


$x^\star$ is optimal

Optimal and Pareto optimal points (cont'd)

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

i.e., $\quad y$ feasible, $\quad f_0(y) \preceq_K f_0(x) \quad \Longrightarrow \quad f_0(x) = f_0(y)$



x^{po} is Pareto optimal

if there are multiple Pareto optimal values, there is a trade-off between the objectives

Example

- Risk return trade-off in portfolio optimization

$$\begin{array}{ll} \text{minimize (w.r.t. } \mathbf{R}^2_+) & (-\bar{p}^T x, x^T \Sigma x) \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0 \end{array}$$

- $x \in \mathbf{R}^n$ is investment portfolio; x_i is fraction invested in asset i
- $p \in \mathbf{R}^n$ is vector of asset return, modeled as a random variable with mean $\bar{p}$, covariance Σ
- $\bar{p}^T x = \mathbf{E} p^T x$ is expected return; $x^T \Sigma x = \mathbf{Var} p^T x$ is return variance

Scalarization

to find Pareto optimal points: choose $\lambda \succ_{K^*} 0$ and solve scalar problem

$$\begin{array}{ll} \text{minimize} & \lambda^T f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

if x is optimal for scalar problem, then it is Pareto-optimal for vector optimization problem

dual cone of a cone K: (Chapter 2.6)

$$K^* = \{y \mid y^T x \ge 0 \text{ for all } x \in K\}$$

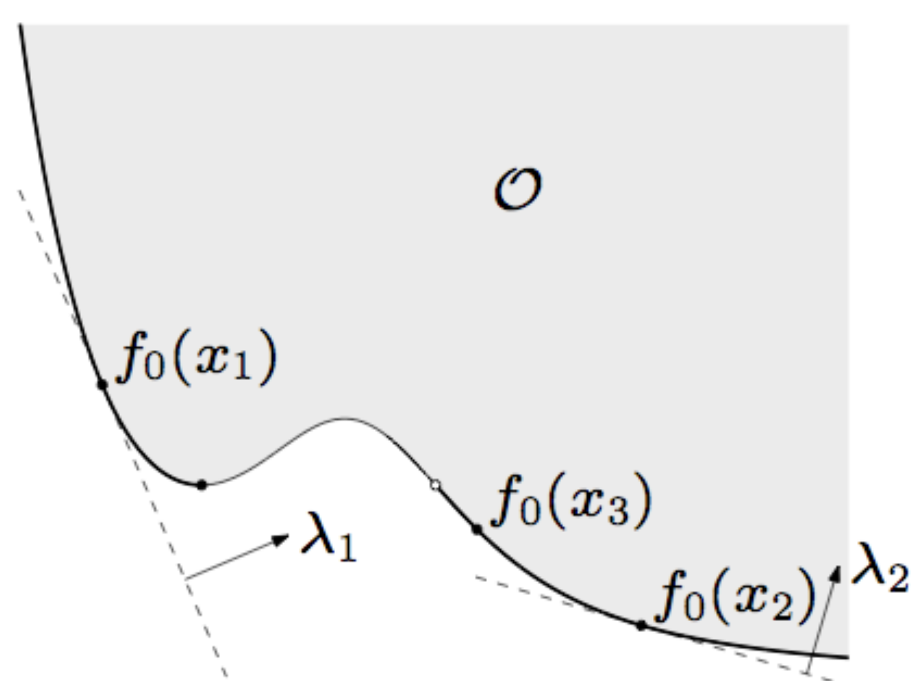
$$y \succeq_{K^*} 0 \quad \Longleftrightarrow \quad y^T x \ge 0 \text{ for all } x \succeq_K 0$$

Scalarization

to find Pareto optimal points: choose $\lambda \succ_{K^*} 0$ and solve scalar problem

$$\begin{array}{ll} \text{minimize} & \lambda^T f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

for convex vector optimization problems, can find (almost) all Pareto optimal points by varying $\lambda \succ_{K^*} 0$



Example

- Risk-return trade-off

$$\begin{array}{ll} \text{minimize} & -\bar{p}^T x + \gamma x^T \Sigma x \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0 \end{array}$$

for fixed $\gamma > 0$, a quadratic program

- Reference
  - Chapter 4.4, 4.5, 4.7, Convex Optimization.

- Acknowledgement
  - Some materials are extracted from the slides created by Prof. Stephen Boyd for the textbook
  - Some materials are extracted from the lecture notes Convex Optimization by Prof. Wei Yu at the University of Toronto