2023.02.02

Convex Optimization Problem

임승현, 윤현석, 김민주

4.1 Optimization Problems

Optimization problem

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

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

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

explicit constraints
$$
\begin{aligned}
f_i(x) &\leq 0 \\
h_i(x) &= 0
\end{aligned}
$$

feasible set: Implicit / explicit constraints 를 만족하는 x의 집합

Optimal value, Optimal point

Optimal value

$$p^\star = \inf\{f_0(x) \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}$$

- $p^\star = \infty$ if problem is infeasible
- $p^\star = -\infty$ if problem is unbounded below

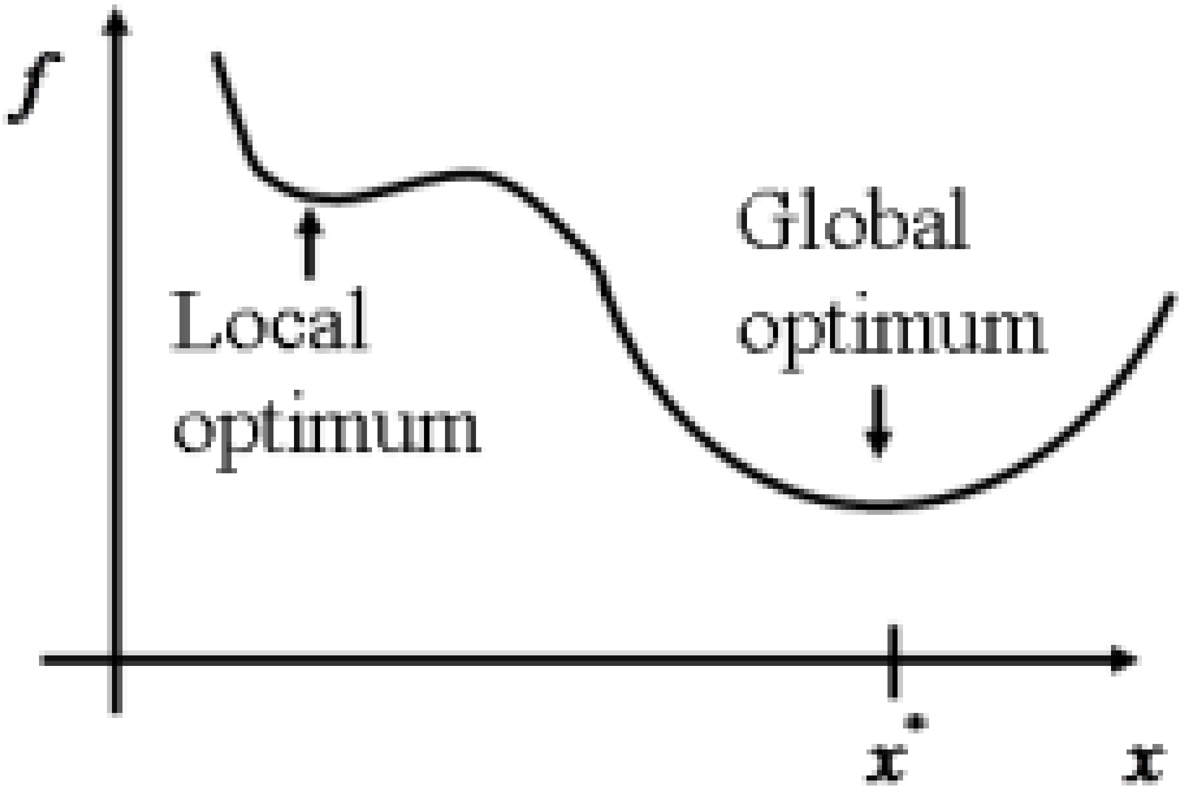
Optimal point

$$X_{\text{opt}} = \{x \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p,\ f_0(x) = p^\star\}$$

Locally optimal

x는 feasible point 이고, $R > 0$ 일 때,

$$f_0(x) = \inf\{f_0(z) \mid f_i(z) \leq 0,\ i = 1, \ldots, m,\\ h_i(z) = 0,\ i = 1, \ldots, p,\ \|z - x\|_2 \leq R\}$$



Examples

- $f_0(x) = x \log x$
- $f_0(x) = 1/x$

4.2 Convex Optimization

Convex optimization problem

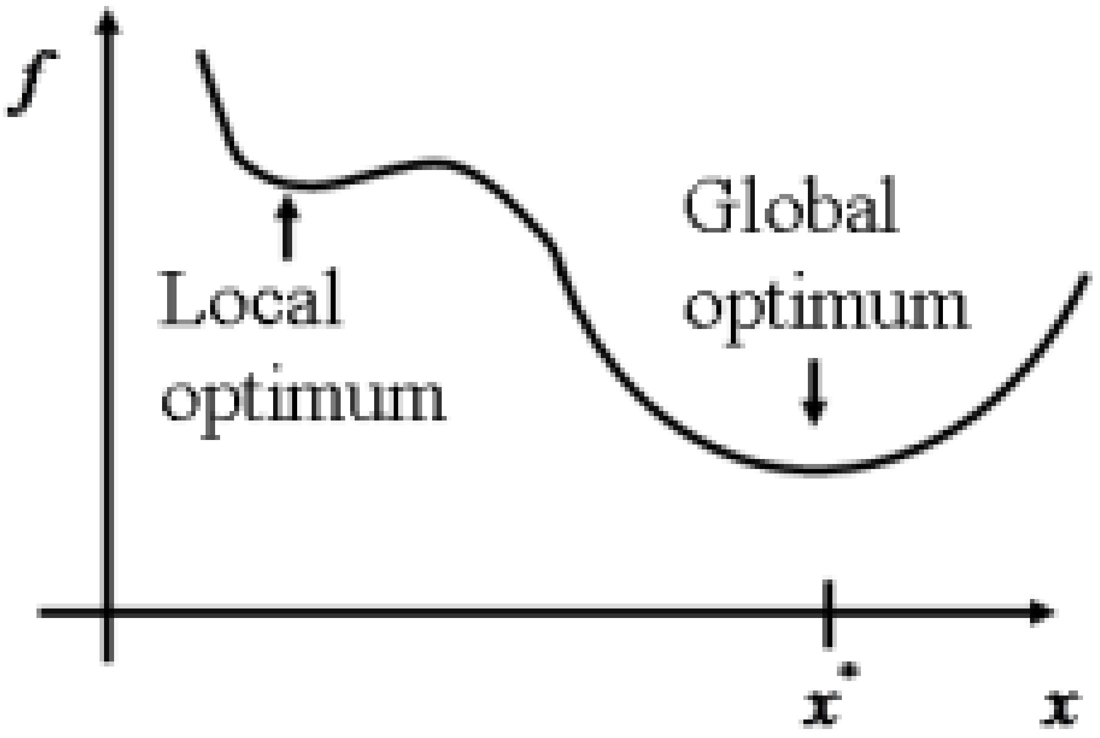
$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \le 0, & i = 1, \ldots, m \\ & a_i^T x = b_i, & i = 1, \ldots, p \end{array} \qquad (Ax = b)$$

- the objective function must be convex,
- the inequality constraint functions must be convex,
- the equality constraint functions $h_i(x) = a_i^T x - b_i$ must be affine.

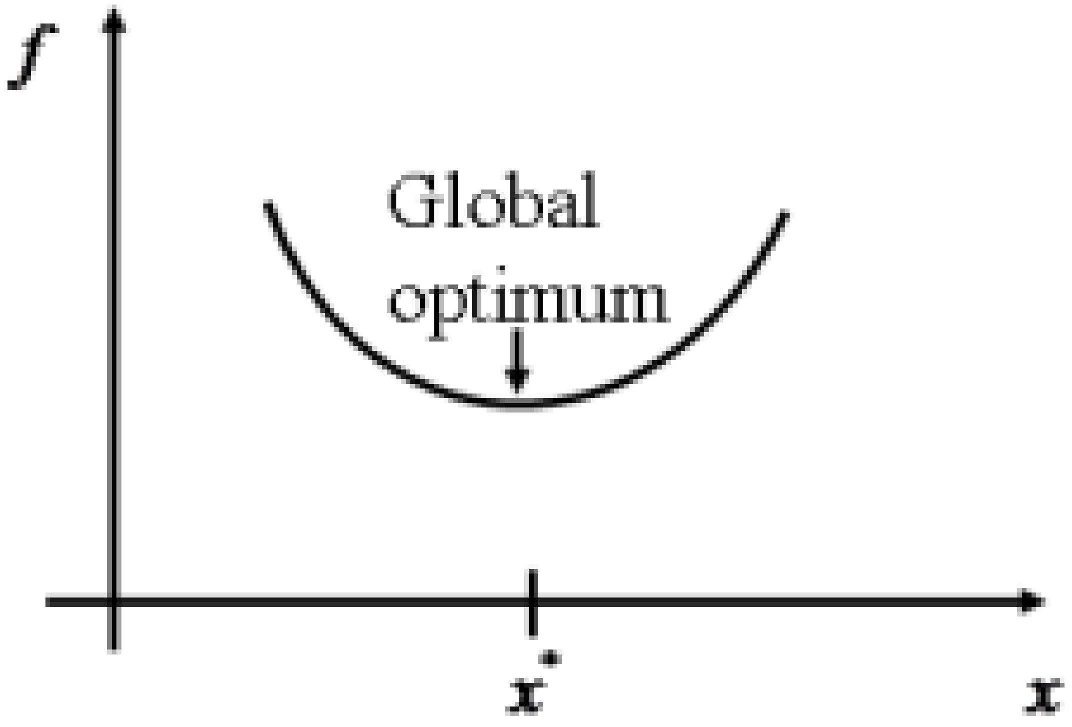
- Convex optimization problem의 feasible set은 convex set!

Local and global optima

Nonconvex



Convex



- Convex optimization problem에서 locally optimal point는 globally optimal point!

Proof

- x를 locally optimal point로 설정

$$f_0(x) = \inf\{f_0(z) \mid z \text{ feasible}, \ \|z - x\|_2 \le R\}$$

- x가 globally optimal 하지 않다고 가정

$$f_0(y) < f_0(x) \qquad \| y - x \|_2 > R$$

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

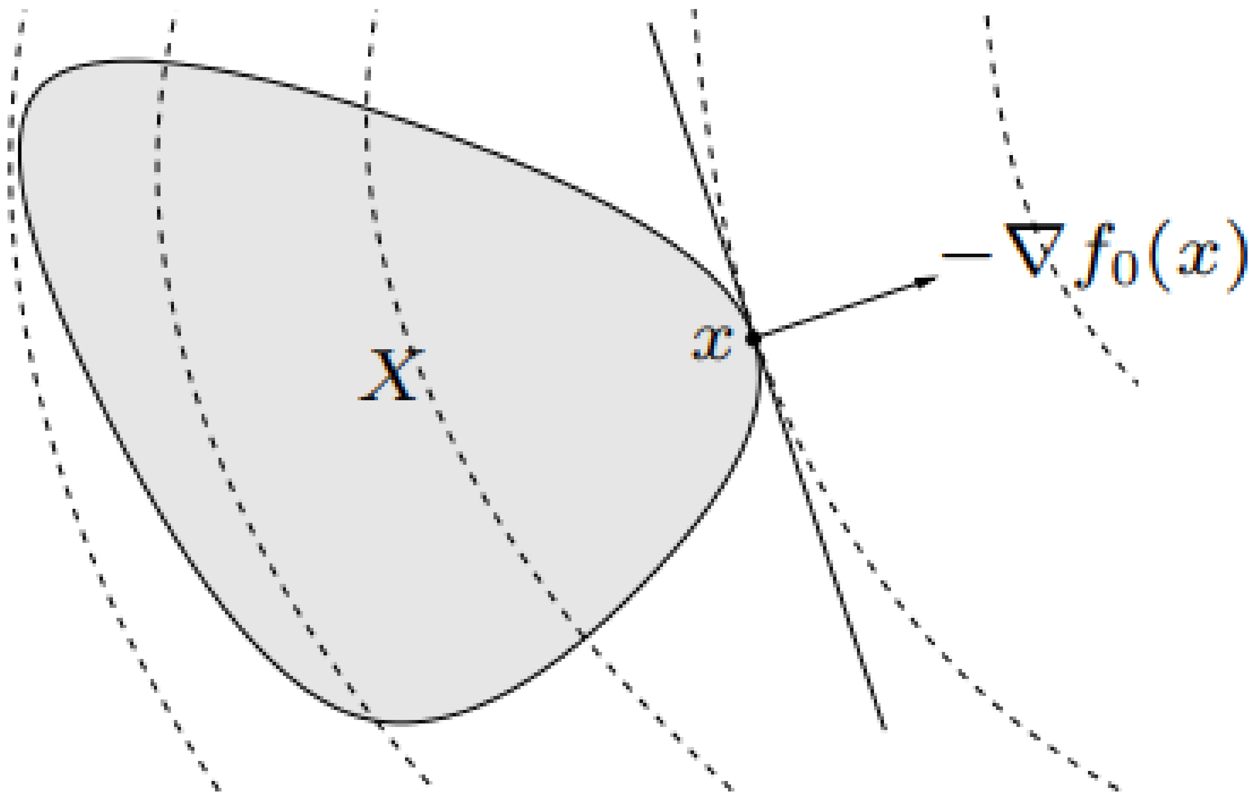
$$\|z - x\|_2 = R/2 < R$$

- 가정에 대한 모순 발생

$$f_0(z) \le \theta f_0(y) + (1 - \theta) f_0(x) < f_0(x)$$

An optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and $\nabla f_0(x)^T(y-x) \geq 0$ for all feasible y



$\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

Proof

- x가 optimal이지만, $\nabla f_0(x)^T(y-x) \geq 0$ 을 만족하지 않는다고 가정

$$\nabla f_0(x)^T(y-x) < 0$$

Consider the point $z(t) = ty + (1-t)x$, where $t \in [0,1]$

$$\left.\frac{d}{dt} f_0(z(t))\right|_{t=0} = \nabla f_0(x)^T(y-x) < 0$$

- 가정에 대한 모순 발생

$$f_0(z(t)) < f_0(x)$$

Equivalent problems

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

- Eliminating equality constraints
- Introducing equality constraints
- Slack variable
- Minimizing over some variables
- Epigraph form

Eliminating/Introducing equality constraints

Eliminating

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

$$\begin{array}{ll} \text{minimize (over } z\text{)} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

$Ax = b \quad \Longleftrightarrow \quad x = Fz + x_0$ for some z

Introducing

$$\begin{array}{ll} \text{minimize} & f_0(A_0 x + b_0) \\ \text{subject to} & f_i(A_i x + b_i) \leq 0, \quad i = 1, \ldots, m \end{array}$$

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

Slack variables, Minimizing over some variables

Slack variables

$$
\begin{array}{ll}
\min_x & f(x) \\
\text{subject to} & g_i(x) \leq 0, i = 1, .., m \\
& Ax = b.
\end{array}
\qquad
\begin{array}{ll}
\min_{x,s} & f(x) \\
\text{subject to} & s_i \geq 0, i = 1, .., m \\
& g_i(x) + s_i = 0, i = 1, \ldots m \\
& Ax = b.
\end{array}
$$

Minimizing over some variables

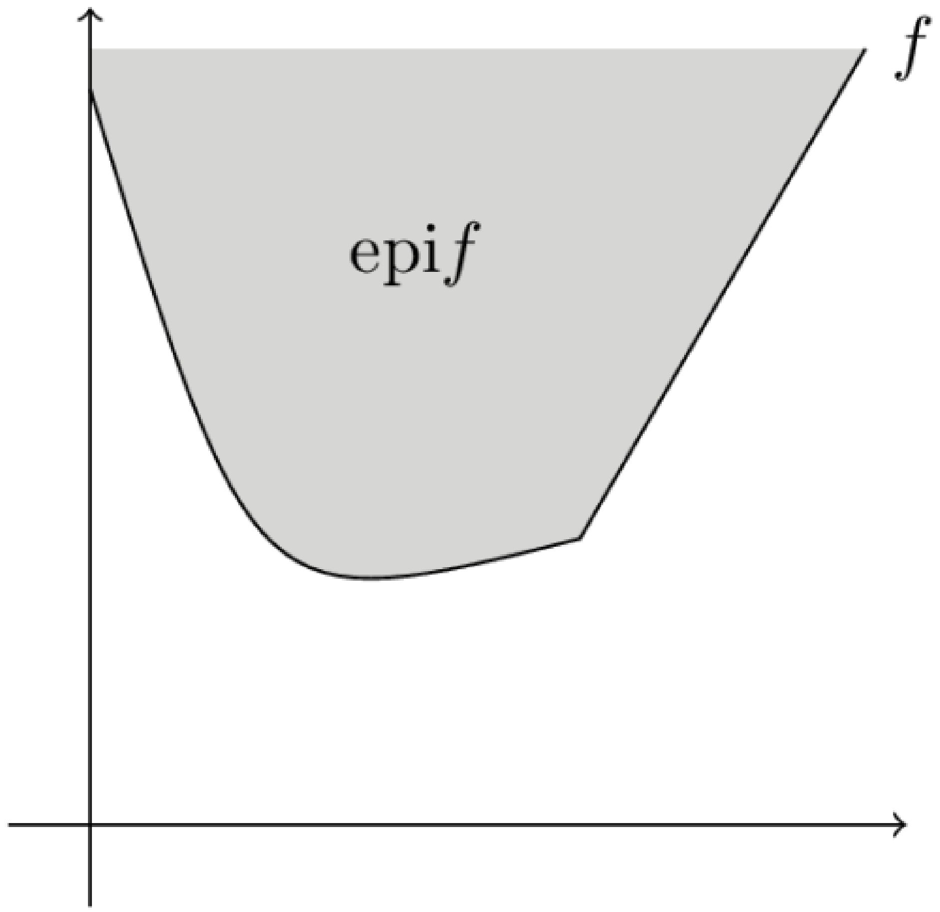
$$
\begin{array}{ll}
\text{minimize} & f_0(x_1, x_2) \\
\text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m
\end{array}
\qquad
\begin{array}{ll}
\text{minimize} & \tilde{f}_0(x_1) \\
\text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m
\end{array}
$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

Epigraph problem form

Epigraph problem form

$$\mathbf{epi}\, f = \{(x,t) \mid x \in \mathbf{dom}\, f,\ f(x) \leq t\}$$



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

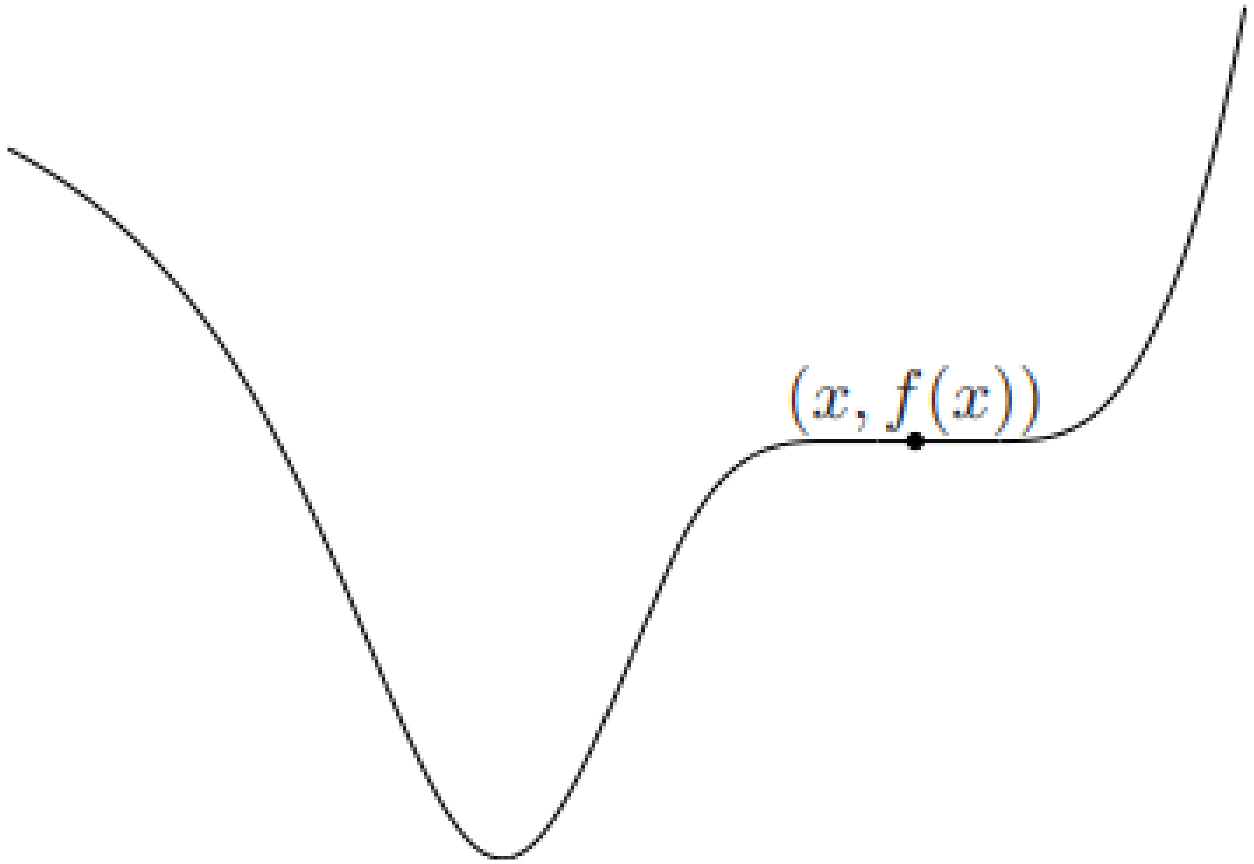
$$\begin{array}{ll} \text{minimize (over } x,\ t) & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

Quasiconvex optimization

Recall

$f : \mathbf{R}^n \to \mathbf{R}$ is quasiconvex if $\mathbf{dom}\, f$ is convex and the sublevel sets

$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ are convex for all α



Quasiconvex optimization problem

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

Locally optimal solution이 globally optimal이 아닐 수도 있다!

Quasiconvex optimization

Recall

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

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

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

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & Ax = b, \end{array} \qquad\qquad \begin{array}{ll} \text{find} & x \\ \text{subject to} & \phi_t(x) \le 0 \\ & f_i(x) \le 0, \quad i = 1, \dots, m \\ & Ax = b, \end{array}$$

if feasible, we can conclude that $t \ge p^\star$

convex feasibility problems

Bisection method for quasiconvex optimization

Bisection method

근을 포함하는 구간을 임의로 정하고, 그 구간을 절반씩 줄여 나가며 근을 구하는 방법

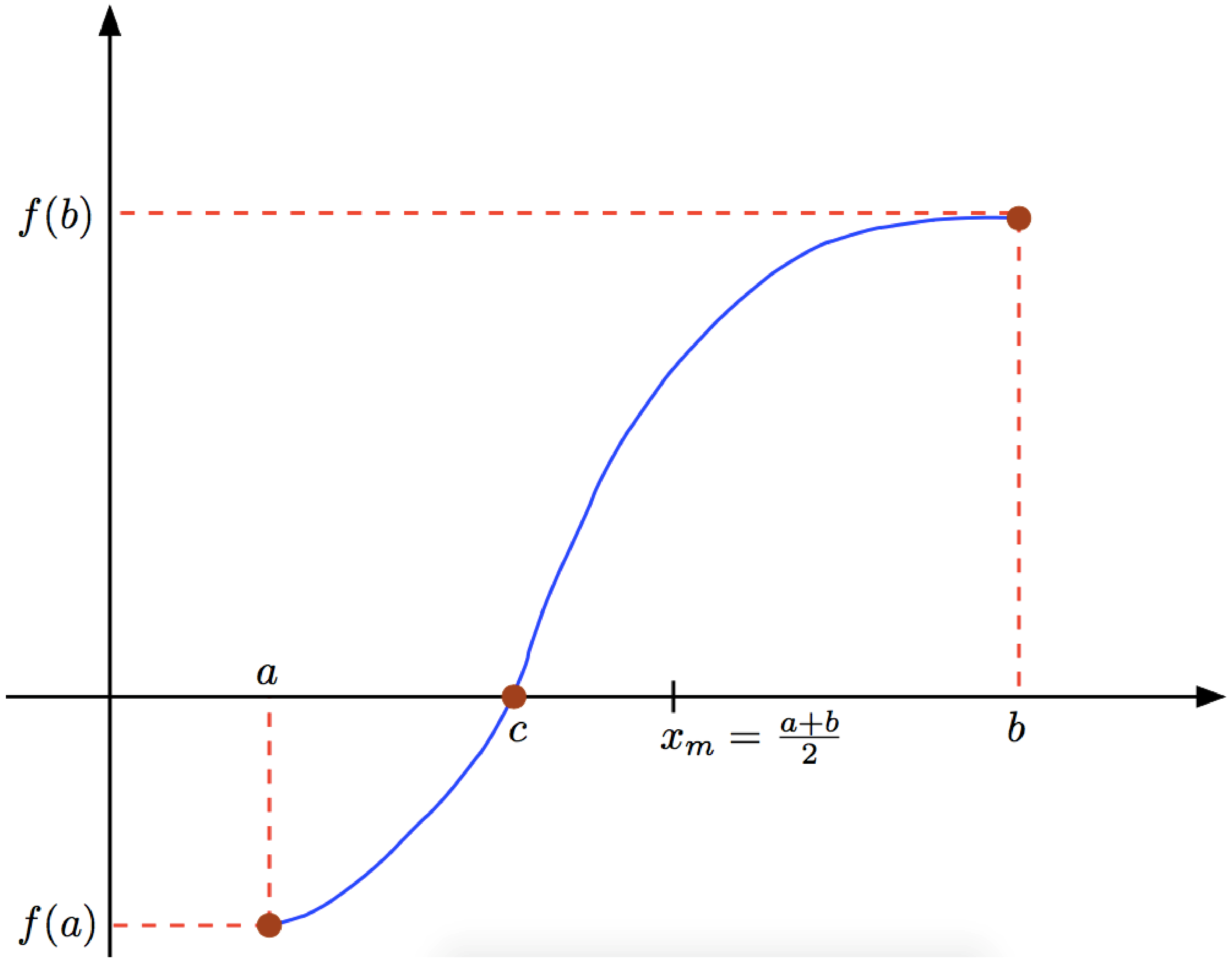
$$
\begin{array}{ll}
\text{find} & x \\
\text{subject to} & \phi_t(x) \le 0 \\
& f_i(x) \le 0, \quad i = 1, \ldots, m \\
& Ax = b,
\end{array}
$$

given $l \le p^\star$, $u \ge p^\star$, tolerance $\epsilon > 0$.

repeat

1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (4.26).
3. **if** (4.26) is feasible, $u := t$; **else** $l := t$.

until $u - l \le \epsilon$.



과제

Exercise 4.1

Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, \quad x_2 \geq 0. \end{array}$$

Make a sketch of the feasible set. For each of the following objective functions, give the optimal set and the optimal value.

(a) $f_0(x_1, x_2) = x_1 + x_2$.

(b) $f_0(x_1, x_2) = -x_1 - x_2$.

(c) $f_0(x_1, x_2) = x_1$.

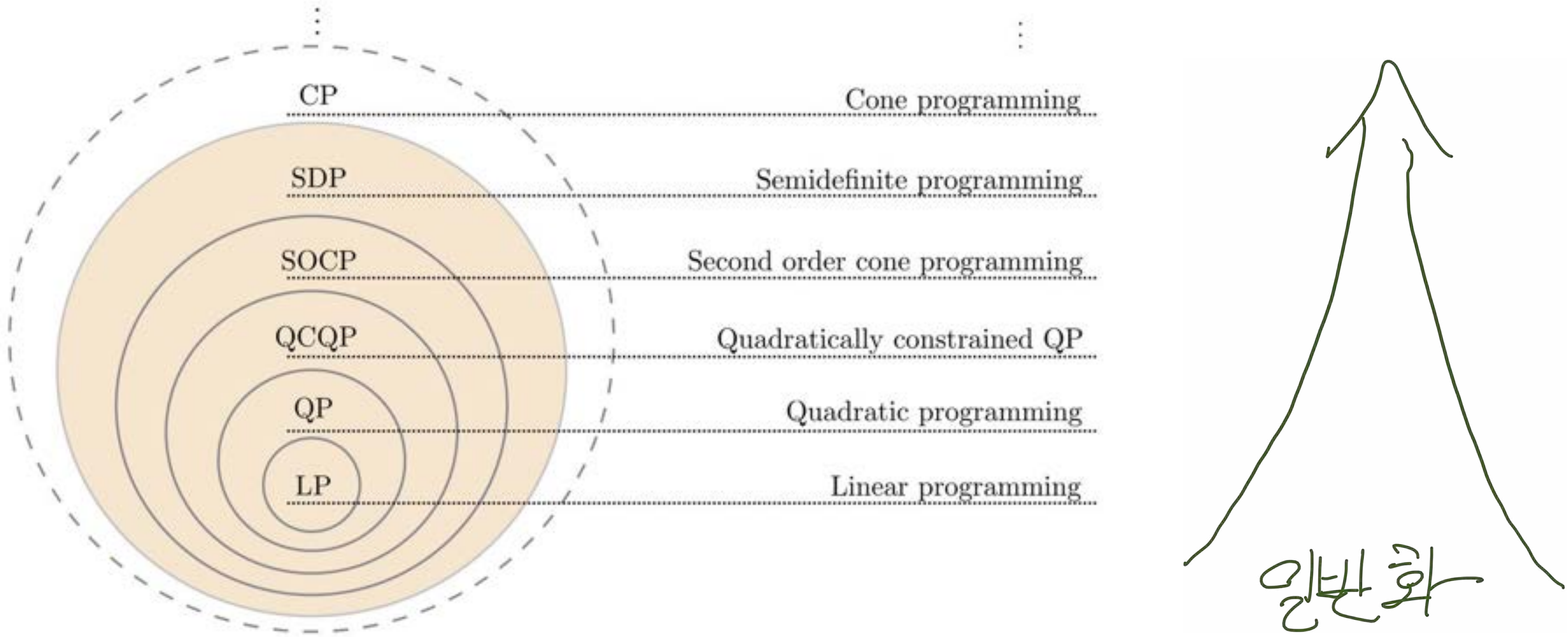
(d) $f_0(x_1, x_2) = \max\{x_1, x_2\}$.

(e) $f_0(x_1, x_2) = x_1^2 + 9x_2^2$.

Convex Optimization problem

ESC 28기 윤현석

포함 관계


CP
SDP
SOCP
QCQP
QP
LP
Cone programming
Semidefinite programming
Second order cone programming
Quadratically constrained QP
Quadratic programming
Linear programming
일반화

LP

Linear optimization Problems

minimize: $c^T x + d$ Affine

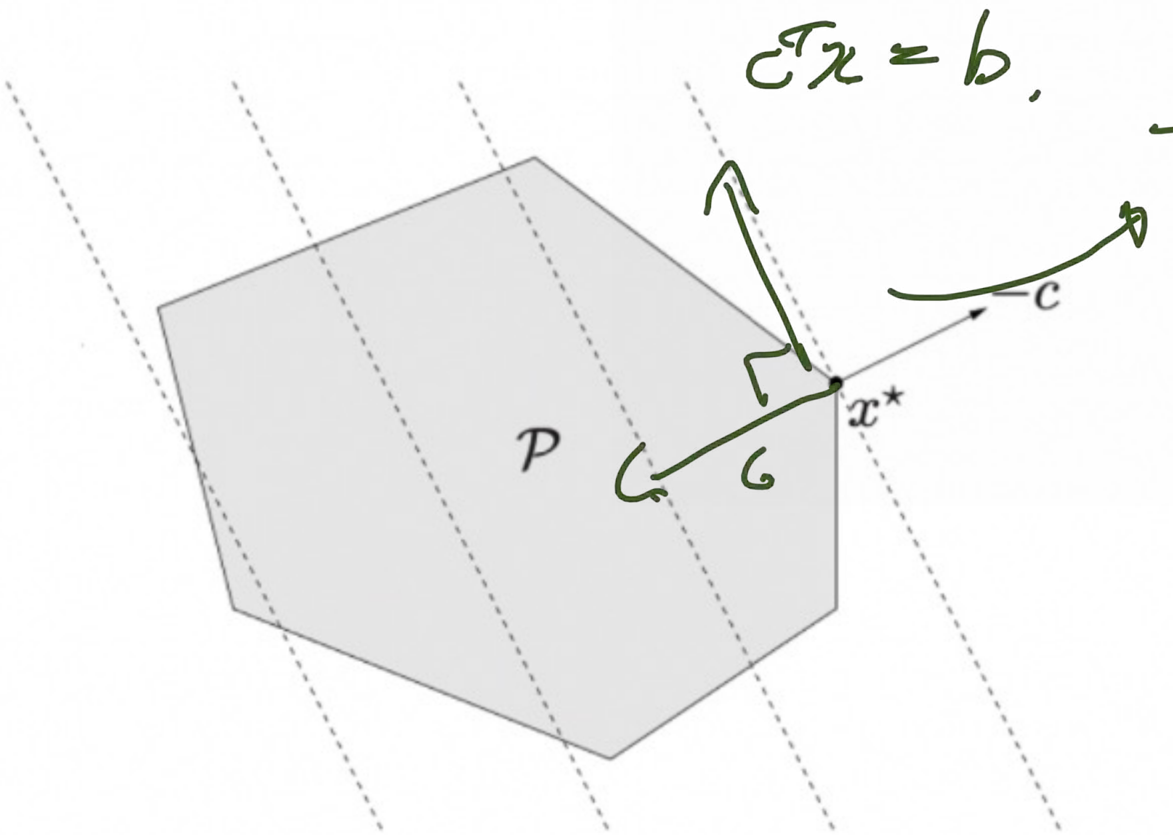
subject to: Affine

$Gx \preceq h$

$Ax = b,$

$where\, G \in \mathbf{R}^{m \times n}\; and\; A \in \mathbf{R}p \times n$

Linear optimization Problems



hyperplane은 c와 직각.
가장 멀어지는 방향으로.

Polyhedron 위에서
최적점 찾기

선이 affine function

Linear optimization Problems

자주 만나는 LP 들

Standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0. \end{array}$$

Inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b. \end{array}$$

Inequality constraint 를 componentwise nonnegativity constraints x ≽ 0 만을 가짐

Equality constraint 가 없음

Linear optimization Problems

general LP 를 standard form 으로 바꾸는 것이 유용할 수 있다.
왜? → standard form 을 푸는 테크닉을 사용할 수 있다.

minimize: $c^T x + d$

subject to:

$Gx \preceq h$

$Ax = b,$

$x = x^+ - x^-$

$$
\begin{array}{ll}
\text{minimize} & c^T x + d \\
\text{subject to} & Gx + s = h \\
& Ax = b \\
& s \succeq 0.
\end{array}
$$

$$
\begin{array}{ll}
\text{minimize} & c^T x^+ - c^T x^- + d \\
\text{subject to} & Gx^+ - Gx^- + s = h \\
& Ax^+ - Ax^- = b \\
& x^+ \succeq 0, \quad x^- \succeq 0, \quad s \succeq 0,
\end{array}
$$

componentwise nonnegativity constraint 만 가짐

Linear optimization Problems

Example – Diet problem

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

$$A = \begin{bmatrix} \text{탄} \cdots \\ \text{단} \cdots \\ \text{지} \cdots \end{bmatrix}$$

C: 가격 $= \begin{bmatrix} \vdots \end{bmatrix}$ 음식 별

A : 식단 별 영양소 함량

b: 최소 영양소 요구량

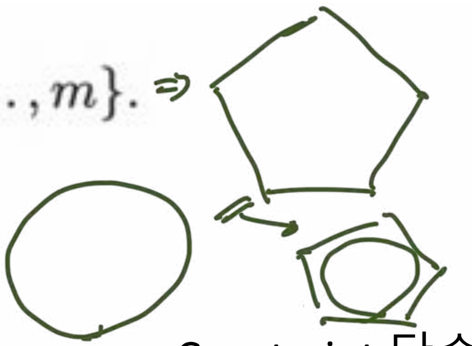
X: 음식의 수량 $= \begin{bmatrix} \vdots \end{bmatrix}$

Linear optimization Problems

Example – Chebyshev center of a polyhedron

Polyhedron $\mathcal{P} = \{x \in \mathbf{R}^n \mid a_i^T x \le b_i,\ i = 1, \ldots, m\}. \Rightarrow$

Optimal ball $\mathcal{B} = \{x_c + u \mid \|u\|_2 \le r\}. \Rightarrow$



With variable $x_c \in R^n$, radius r

Subject to the constraint $\mathcal{B} \subseteq \mathcal{P}$.

$a_i^T u = \|a_i\| \times \|u\| \times \cos\theta$

$\Rightarrow \{ \|u\|_2 \le r,\ \cos\theta \le 1$

Constraint 단순화

$\|u\|_2 \le r \Longrightarrow a_i^T(x_c + u) \le b_i$

$sup\{a_i^T u \mid \|u\|_2 \le r\} = r\|a_i\|_2$ 이므로,

$a_i^T x_c + r\|a_i\|_2 \le b_i$

이 된다.

Linear optimization Problems

Example – Chebyshev center of a polyhedron

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \leq b_i, \quad i = 1, \ldots, m, \end{array}$$

(with variables r and x_c)

Linear optimization Problems

Example – Chebyshev inequalities

우리는 discrete random 변수 x 가 $\{u_1, ..., u_n\} \subseteq \mathbf{R}$ 의 형태로 가정되었을 때 그 확률분포를 고려할 수 있다. 우리는 x 의 분포를 $p \in \mathbf{R}^n$ 을 통해서 나타낼 수 있다.

$p_i = \mathbf{prob}(x = u_i)$ 이렇게. 따라서 p 는 $p \succeq 0, \mathbf{1}^T p = 1$ 이다.

x가 $\{u_1, u_2 \cdots u_n\}$ 값 가질수 있음, with $\{p_1, p_2 \cdots p_n\}$ 의 확률로.

probability distribution

만약 f가 x 의 임의의 함수라면, ⇒ r. v. 의 정의

$\mathbf{E}f = \sum_{i=1}^{n} p_i f(u_i)$ 이다.

$\{u_1 \cdots u_n\}$ - sample space with prob measure ⇒ real valued function 으로 변환시켜 준다.

$$\alpha_i \leq a_i^T p \leq \beta_i, \quad i = 1, \ldots, m.$$

확률의 upper, lower bound 를 정해줄 수 있는 선행지식이 존재한다. -> constraint 역할

chebyshev, markov inequality 등.

Linear optimization Problems

Example – Chebyshev inequalities

$a_0^T = [\cdot\ \cdot\ \cdot\ \cdot\ \cdot]$ r.v $\begin{bmatrix} p \\ \\ \end{bmatrix}$ = 어떤 r.v의 expectation.

$$
\begin{array}{ll}
\text{minimize} & a_0^T p \\
\text{subject to} & p \succeq 0, \quad \mathbf{1}^T p = 1 \\
& \alpha_i \le a_i^T p \le \beta_i, \quad i = 1, \ldots, m,
\end{array}
$$

수통에서 배운 체비셰프 부등식, 마르코프 부등식 처럼 우리가
이미 알고 있는 선행 지식을 이용한 constraint

$$\mathbf{E} f_0(x) = a_0^T p$$

의 Lower and upper bounds 를
찾는 문제.

$LP \subseteq QP$

Quadratic optimization Problems

objective function 이 quadratic,
constraint function 이 affine 인 convex optimization 문제.

$$\begin{array}{ll} \text{minimize} & (1/2)x^T Px + q^T x + r \Rightarrow \text{quadratic form} \\ \text{subject to} & \left.\begin{array}{l} Gx \preceq h \\ Ax = b, \end{array}\right\} = \text{Affine} \end{array}$$

where $P \in \mathbf{S}_+^n$, $G \in \mathbf{R}^{m \times n}$, $A \in \mathbf{R}^{p \times n}$.

Quadratic optimization Problems

포함 관계 $LP \subseteq QP$

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

where $P \in \mathbf{S}_{+}^{n}$, $G \in \mathbf{R}^{m \times n}$, $A \in \mathbf{R}^{p \times n}$.

minimize: $c^T x + d$

subject to:

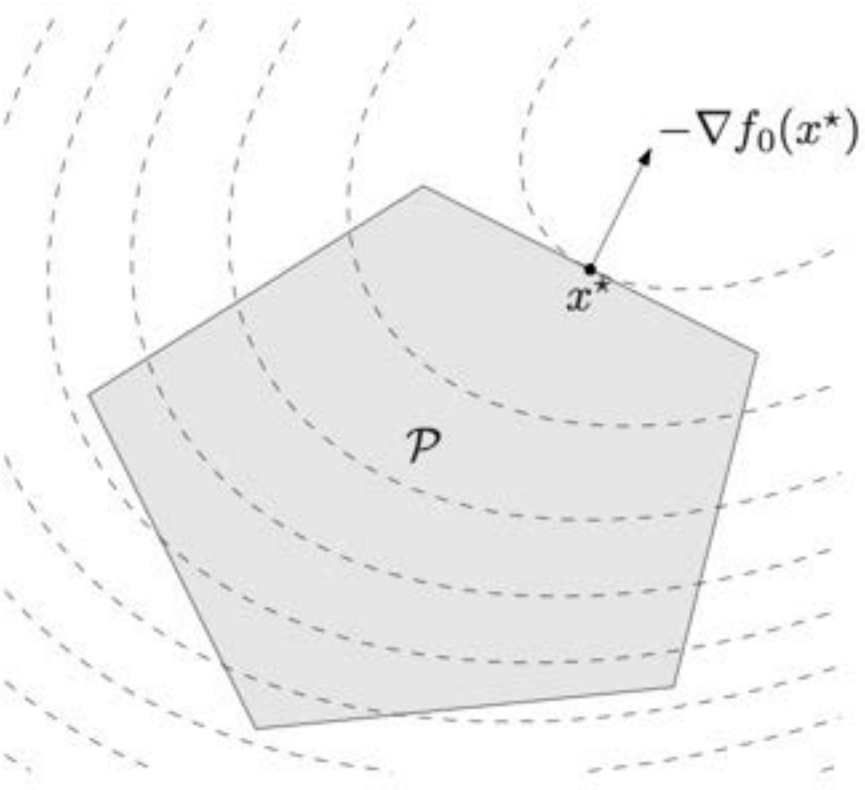
$Gx \preceq h$

$Ax = b,$

$where\ G \in \mathbf{R}^{m \times n}\ and\ A \in \mathbf{R}p \times n$

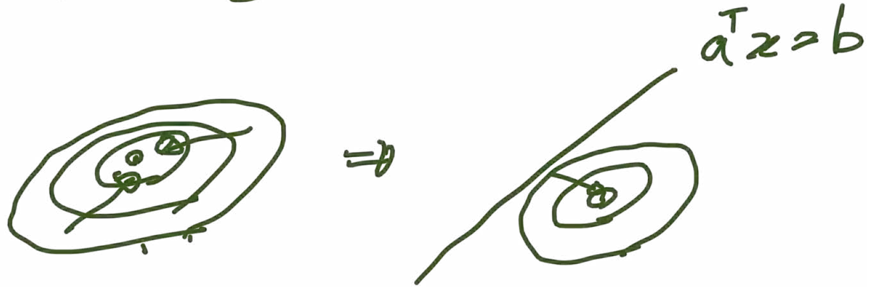
P=0 으로

Quadratic optimization Problems



convex quadratic function 을 polyhedron 위에서
최소화하는 문제.

ex) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = z$ 최소화.



QCQP

Quadratically Constrained Quadratic Program

만약에 convex 최적화 문제의 기본형에서 objective function와 inequality constraint 가 모두 quadratic 이면, 우리는 quadratically constrained quadratic program (QCQP) 라고 부른다.

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

$$\text{where } P_i \in \mathbf{S}_+^n, \; i = 0, 1 \ldots, m$$

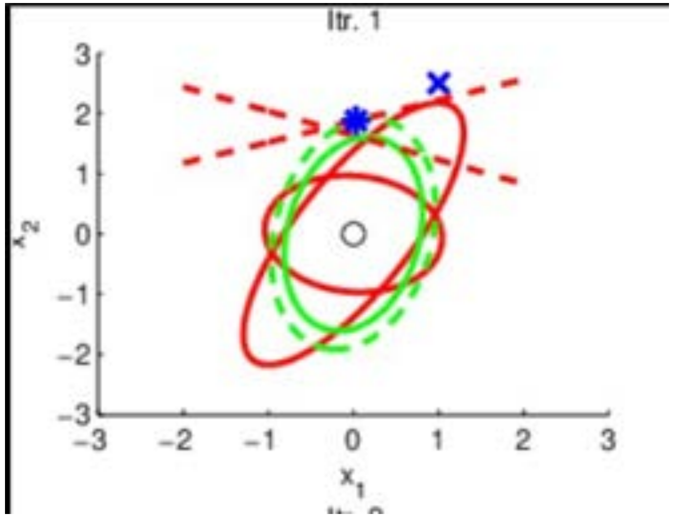
convex quadratic function over a feasible region that is the intersection of ellipsoids ($when\ P_i \succ 0$)

겹쳐진 타원면을 통해서 정의된 feasible region

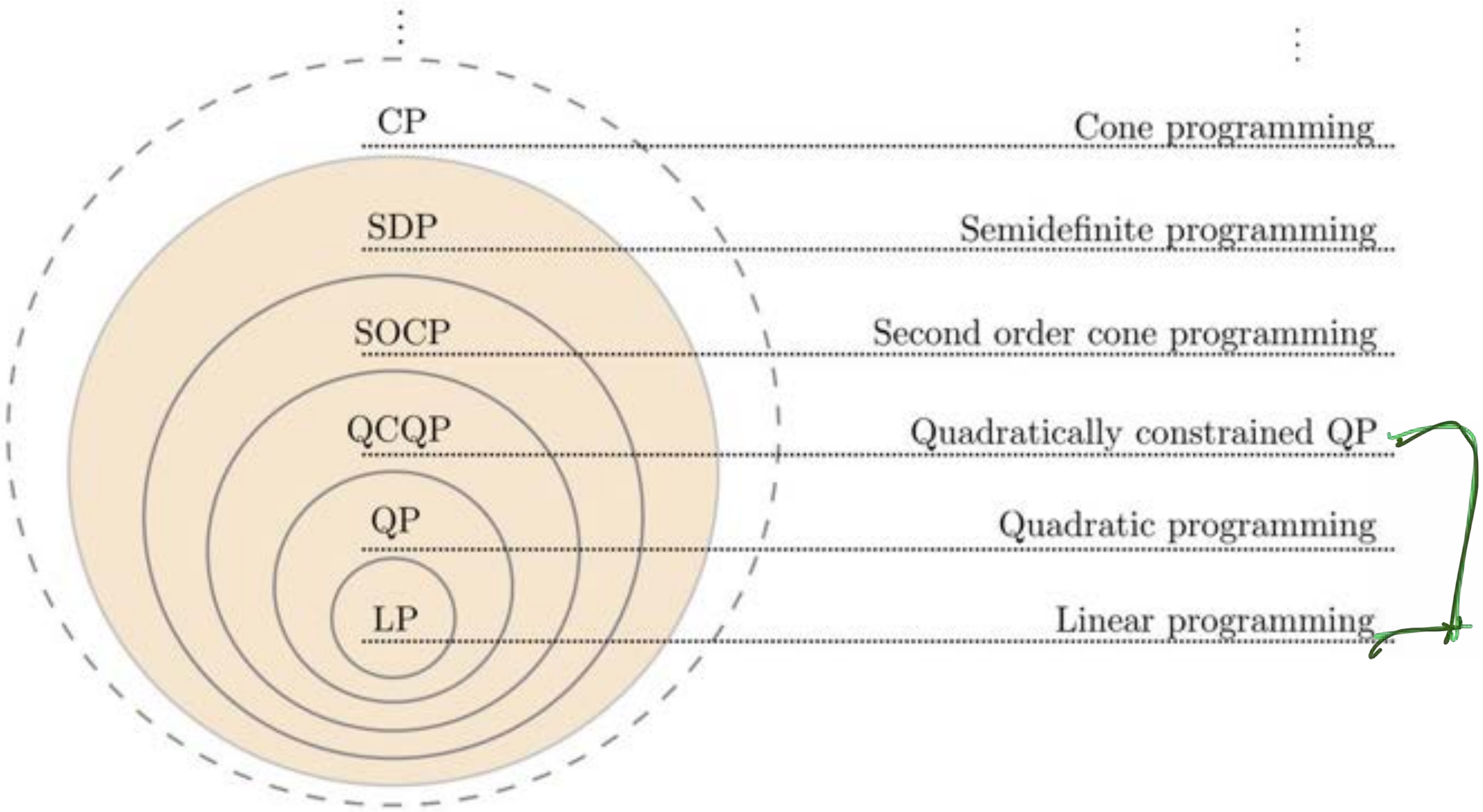
Quadratically Constrained Quadratic Program

convex quadratic function over a feasible region that is the intersection of ellipsoids ($when\ P_i \succ 0$)

겹쳐진 타원면을 통해서 정의된 feasible region



포함 관계


CP
SDP
SOCP
QCQP
QP
LP
Cone programming
Semidefinite programming
Second order cone programming
Quadratically constrained QP
Quadratic programming
Linear programming

포함 관계

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

where $P_i \in \mathbf{S}_+^n,\ i = 0, 1..., m$

$QP \subseteq QCQP$

P_i=0 으로

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

where $P \in \mathbf{S}_+^n,\ G \in \mathbf{R}^{m \times n},\ A \in \mathbf{R}^{p \times n}$.

P=0 으로

minimize: $c^T x + d$

subject to:

$Gx \preceq h$

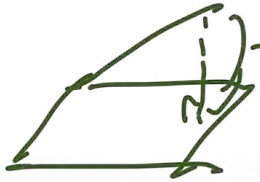
$Ax = b,$

$where\ G \in \mathbf{R}^{m \times n}\ and\ A \in \mathbf{R}p \times n$

$LP \subseteq QP$

Quadratic optimization Problems

Example – Least-squares and regression



→ 최소화

$$
\begin{array}{ll}
\text{minimize} & \|Ax - b\|_2^2 \\
\text{subject to} & l_i \leq x_i \leq u_i, \quad i = 1, \ldots, n,
\end{array}
$$

$$\|Ax - b\|_2^2 = x^T A^T A x - 2b^T A x + b^T b$$ → quadratic.

이므로, QP 이다.

Quadratic optimization Problems

Example – Bounding variance

unknown 분포의 랜덤변수의 분산

n=2이면,

$f^2p - f^2p^2 \Rightarrow$

con.

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

이고, concave quadradic function of p 이다. 따라서 maximization 문제가 가능.

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^n f_i^2 p_i - \left(\sum_{i=1}^n f_i p_i\right)^2 \\ \text{subject to} & p \succeq 0, \quad \mathbf{1}^T p = 1 \\ & \alpha_i \le a_i^T p \le \beta_i, \quad i = 1, \ldots, m. \end{array}$$

probability distribution

선행지식

Maximum possible variance 구하기

Quadratic optimization Problems

Example – Markowitz portfolio optimization

$r = p^T x$: overall return

x_i ... ; position. p_i : price
with 관련성

constraint

$x_i \geq 0$ (no short position)

$\mathbf{1}^T x = B$ (전체 예산이 B)

stochastic model for price changes

$p \in \mathbf{R}^n$ - random vector: $mean : \bar{p}$ $covariance : \Sigma$

따라서 return 또한: $mean : \bar{p}^T x$ $variance : x^T \Sigma x$

Quadratic optimization Problems

Example – Markowitz portfolio optimization

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

안정적인 포트폴리오를 구성하기 위한 문제 with 최소 수익률

Quadratic optimization Problems

Example – Markowitz portfolio optimization

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

확장 – 숏 가즈아

$$x_{long} \succeq 0, \quad x_{short} \succeq 0, \quad x = x_{long} - x_{short}, \quad \underbrace{\mathbf{1}^T x_{short} \leq \eta \mathbf{1}^T x_{long}}_{\text{숏 제외}}$$

확장 – 유동성 없음

$$x = x_{init} + u_{buy} - u_{sell}, \quad u_{buy} \succeq 0, \quad u_{sell} \succeq 0$$

산돈 판돈

$$(1 - f_{sell}) \mathbf{1}^T u_{sell} = (1 + f_{buy}) \mathbf{1}^T u_{buy}$$

현금X.

$$f_{buy} \geq 0, f_{sell} \geq 0$$

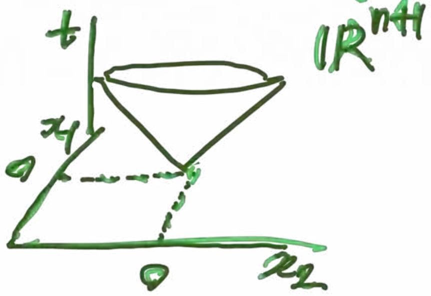
(편의를 위해 모든 자산에 대해 같다고 설정)
(팔때는 돈이 들어오는 것에서 빼주고, 살때는 돈이 나가는 것에서 더해주는 논리)
(초기에 산 것 + 이후 산 것 + 판 것 = x = 현재 포트폴리오)
(그러면 위의 조건은 zero net cash - 자산 중 현금을 남겨놓지 않겠다는 것.)

SOCP

Second-Order Cone Programming

$\{(A_i x + b_i,\ c_i^T x + d_i) \mid \ \} \in \mathbb{R}^{n+1}$

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



$(x_1, x_2) \in \mathbb{R}^{n+1}$

Constraint 의 형태

$$C = \{(x,t) \mid \|x\| \leq t\} \subseteq \mathbf{R}^{n+1}$$

Icecream cone 의 형태와 같다.

$c_i = 0, i = 1, \ldots, m$ 일 때, SOCP 는 QCQP 와 equivalent 해진다.

$$LP \subseteq QP \subseteq QCQP \subseteq SOCP$$

라고 하지만,,,,

Second-Order Cone Programming

조금 더 와닿는 유도 과정

https://convex-optimization-for-all.github.io/contents/chapter05/2021/02/08/05_04_Second_Order_Cone_Programming_(SOCP)/

QCQP ⊆ SOCP

※ 참고

또한 $A_i = 0, i = 1, ..., m$, 이면 SOCP 는 LP 로 축약된다.

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

$\rightarrow \|b_i\|_2 \le c_i^T x + d_i$

$-c_i^T x \le d_i - \|b_i\|_2$

$\vdots$

minimize: $c^T x + d$

subject to:

$Gx \preceq h$

$Ax = b,$

where $G \in \mathbf{R}^{m \times n}$ *and* $A \in \mathbf{R}p \times n$

Second-Order Cone Programming

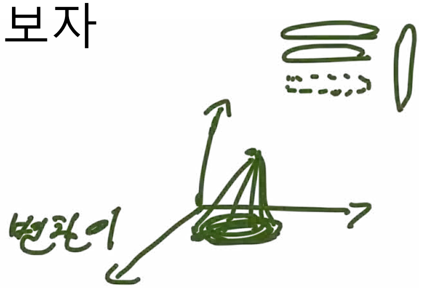
Example - Robust linear programming

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

인자들이 랜덤한 경우. 예를 들어 a_i 가 랜덤한 경우를 생각해 보자

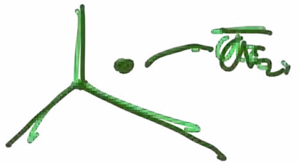
$$a_i \in \mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\} \Rightarrow$$

where $P_i \in \mathbf{R}^{n \times n}$



(만약 P_i 가 비가역이라면 우리는 rank P_i의 차원에서 'flat'한 ellipsoid 를 얻게 된다. full rank 를 ellipsoid 가 가지지 못한다는 것이고 그러면 variance 또한 full dimension 에서 놀지 못하고 최소 한 차원 낮게 놀게 된다는 것.

P_i=0는 a_i 가 완전하게 known 임을 의미한다.)



우리는 a_i의 모든 가능한 값에 대해 만족하는 constraint 를 원한다. = robust linear program.

Second-Order Cone Programming

Example - Robust linear programming

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

Constraint 를 다시 쓰자면

$$(\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + u^T P_i^T x$$

$$= \bar{a}_i^T x + \|u^T\| \times \|P_i^T x\| \times \cos\theta$$

$$sup\{a_i^T x | a_i \in \mathcal{E}_i\} \le b_i$$

$$\begin{aligned} sup\{a_i^T x | a_i \in \mathcal{E}_i\} &= \bar{a}_i^T x + sup\{u^T P_i^T x | \, ||u||_2 \le 1\} \\ &= \bar{a}_i^T x + ||P_i^T x||_2 \end{aligned}$$

utu = 1

$$1 \times \|P_i^T x\|$$

$$\bar{a}_i^T x + ||P_i^T x||_2 \le b_i$$

Second-order cone constraint $||P_i^T x||_2 \le b_i - \bar{a}_i^T x$

Second-Order Cone Programming

Example - Robust linear programming

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

Second-Order Cone Programming

Example - Linear programming with random constraints

Robust LP 를 통계적인 framework 로 표현해 보자

$a_i = gaussian\ random\ vector$ 정규분포.

$mean: \bar{a}_i$

$variance: \Sigma_i$

$\eta \geq \frac{1}{2}$

$\mathbf{prob}(a_i^T x \leq b_i) \geq \eta$

$u = a_i^T x,\ \sigma^2$ 이라고 할 때,

$\mathbf{prob}\left(\frac{u-\bar{u}}{\sigma} \leq \frac{b_i-\bar{u}}{\sigma}\right) \geq \eta$

cdf – monotone.

$\Phi\left(\frac{b_i-\bar{u}}{\sigma}\right) \geq \eta$

$\Phi\left(\frac{b_i-\bar{u}}{\sigma}\right)$ - cdf

따라서 $\mathbf{prob}\left(\frac{u-\bar{u}}{\sigma} \leq \frac{b_i-\bar{u}}{\sigma}\right) \geq \eta$ 은 $\frac{b_i-\bar{u}}{\sigma} \geq \Phi^{-1}(\eta)$ → $\bar{u} + \Phi^{-1}(\eta)\sigma \leq b_i$

$\bar{u} = \bar{a}_i^T x$ 이고, $\sigma = (x^T \Sigma_i x)^{1/2}$ 이므로, 우리는 위의 식을

$\bar{a}_i^T x + \Phi^{-1}(\eta)||\Sigma_i^{1/2} x||_2 \leq b_i$ 으로 쓸 수 있다. → $\Phi^{-1}(\eta)||\Sigma_i^{1/2} x||_2 \leq b_i - \bar{a}_i^T x$

SOCP의 형태

Second-Order Cone Programming

Example - Linear programming with random constraints

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

$$\Downarrow$$

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

HW

4.29 *Maximizing probability of satisfying a linear inequality.* Let c be a random variable in $\mathbf{R}^n$, normally distributed with mean $\bar{c}$ and covariance matrix R. Consider the problem

$$\begin{array}{ll} \text{maximize} & \mathbf{prob}(c^T x \geq \alpha) \\ \text{subject to} & Fx \preceq g, \quad Ax = b. \end{array}$$

Assuming there exists a feasible point $\tilde{x}$ for which $\bar{c}^T \tilde{x} \geq \alpha$, show that this problem is equivalent to a convex or quasiconvex optimization problem. Formulate the problem as a QP, QCQP, or SOCP (if the problem is convex), or explain how you can solve it by solving a sequence of QP, QCQP, or SOCP feasibility problems (if the problem is quasiconvex).

Convex optimization problems

- 4.6 Generalized inequality constraints
- 4.7 Vector optimization

임승현 윤현석 김민주

01

02

Generalized inequality constraints

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \ldots, m \\
& Ax = b,
\end{array}
$$

Conic form problem

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Fx + g \preceq_K 0 \\
& Ax = b.
\end{array}
$$

Semidefinite program(SDP)

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

$G,\ F_1, \ldots, F_n \in \mathbf{S}^k$, and $A \in \mathbf{R}^{p \times n}$.

Multiple independent LMI는 single LMI로 변환

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \leq 0,\ x_1 \bar{F}_1 + \cdots + x_n \bar{F}_n + \bar{G} \leq 0$$

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \bar{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \bar{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \bar{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \bar{G} \end{bmatrix} \leq 0.$$

SDP as generalization

LP as SDP

LP

minimize $c^T x$

subject to $Ax \le b$

»

SDP

minimize $c^T x$

subject to $diag(Ax - b) \le 0$

SDP as generalization

SOCP as SDP

SOCP

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

SDP

$$\begin{array}{ll}\text{minimize} & f^T x\\ \text{subject to} & \begin{bmatrix}(c_i^T x + d_i)I & A_i x + b_i\\ (A_i x + b_i)^T & c_i^T x + d_i\end{bmatrix} \ge 0,\ i = 1,\dots,m\end{array}$$

$$c_i^T x + d_i - \frac{(A_i^T x + b_i)^T (A_i x + b_i)}{c_i^T x + d_i} \ge 0$$

$$(A_i^T x + b_i)^T (A_i x + b_i) \ge (c_i^T x + d_i)^2$$

Schur complement

SDP as generalization

SOCP as SDP

<Schur complement>

Definition:

Let $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in R^{(p+q)\times(p+q)}$.

Then if A is invertible, the **Schur complement S** of the block A in matrix X is defined as $S = C - B^T A^{-1} B$

<Schur complement lemma>

Suppose $A \in R^{p\times p} > 0$ and C is symmetric. Then,

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0 \Leftrightarrow S = C - B^T A^{-1} B \geq 0$$

Fastest mixing Markov chain on a graph

Markov chain

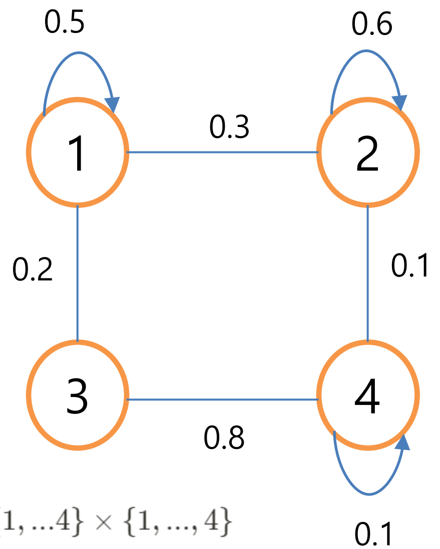
state $X(t) \in \{1, ..., n\}$, $t \in \mathbf{Z}_+$

transition probability matrix P

$$P_{ij} = \mathbf{prob}(X(t+1) = i | X(t) = j), i, j = 1, ..., n$$

$$P_{ij} \geq 0, i, j = 1, ...n \ , \qquad \mathbf{1}^T P = \mathbf{1}^T, \qquad P = P^T$$

$$P\mathbf{1} = \mathbf{1}$$



$$P = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 0.5 & 0.3 & 0.2 & 0 \\ 0.3 & 0.6 & 0 & 0.1 \\ 0.2 & 0 & 0 & 0.8 \\ 0 & 0.1 & 0.8 & 0.1 \end{bmatrix} \end{array}$$

$\epsilon \subseteq \{1, ...4\} \times \{1, ..., 4\}$

$(i, j) \in \epsilon \Leftrightarrow (j, i) \in \epsilon$

$(i, j) \notin \epsilon$ 이면 $P_{ij} = 0$

Fastest mixing Markov chain on a graph

Markov chain

state $X(t) \in \{1, ..., n\}$, $t \in \mathbf{Z}_+$

transition probability matrix P

$P_{ij} = \mathbf{prob}(X(t+1) = i | X(t) = j), i, j = 1, ..., n$

$P_{ij} \geq 0, i, j = 1, ...n$, $\mathbf{1}^T P = \mathbf{1}^T$, $P = P^T$

$P\mathbf{1} = \mathbf{1}$

Equilibrium distribution: $(1/n)\mathbf{1}$ (uniform dist)

$\pi^T P = \pi^T$: chain stops here(stationary)

mixing rate $r = max_{i=2,...,n} |\lambda_i(P)| = max\{\lambda_2, -\lambda_n\}$

$1 = \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$

$$
\begin{aligned}
r &= ||QPQ||_2 \\
&= ||(I - (1/n)\mathbf{1}\mathbf{1}^T)P(I - (1/n)\mathbf{1}\mathbf{1}^T)||_2 \\
&= ||P - (1/n)\mathbf{1}\mathbf{1}^T||_2
\end{aligned}
$$

$Q = I - (1/n)\mathbf{1}\mathbf{1}^T$

minimize $||P - (1/n)\mathbf{1}\mathbf{1}^T||_2$

subject to $P\mathbf{1} = \mathbf{1}$

$P_{ij} \geq 0, i, j = 1, ...n$

$P_{ij} = 0$ for $(i, j) \notin \epsilon$

minimize t

subject to $-tI \preceq P - (1/n)\mathbf{1}\mathbf{1}^T \preceq tI$

$P_{ij} \geq 0, i, j = 1, ...n$

$P_{ij} = 0$ for $(i, j) \notin \epsilon$

Vector optimization

minimize (with respect to K) $f_0(x)$

subject to $f_i(x) \leq 0, i = 1, ..., m$

$h_i(x) = 0, i = 1, ..., p$

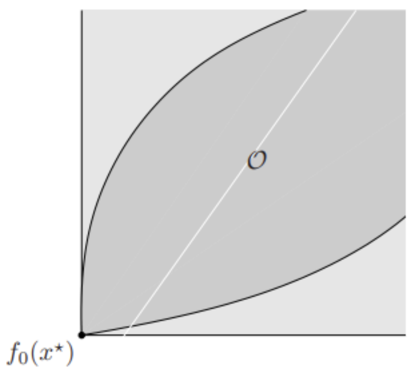
Optimal and Pareto optimal points

optimal

Optimal: $f_0(x)$가 minimum value

minimum element가 있다는 것은

어떤 점 x^*가 모든 feasible y에 대해 $f_0(x^*) \preceq_K f_0(y)$

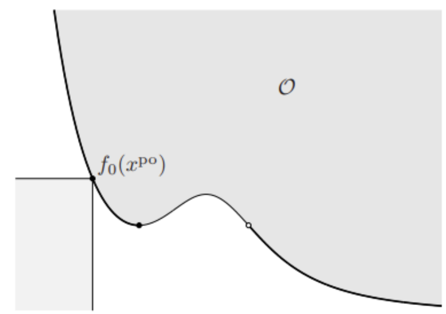


$$\mathcal{O} \subseteq x^* + K$$

Pareto optimal

Pareto optimal: $f_0(x)$가 minimal value

any feasible y에 대해 $f_0(y) \preceq_K f_0(x)$면 $f_0(y) = f_0(x)$



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

Scalarization

Proper dual cone으로 generalized inequality를 정의해보면 다음과 같다. 어떤 점 y 가 있을 때 K 의 모든 점 x 와 내적을 해서 0보다 크다면, y 는 dual cone K^* 에서 0보다 크다.

이때, $\succeq_{K^*}$ 를 $\succeq_K$ 의 **dual**이라고 한다. 즉, **dual generalized inequality**이다.

> $y \succeq_{K^*} 0 \iff y^T x \geq 0$ for all $x \succeq_K 0$

Generalized inequality와 dual의 주요 속성

- $x \preceq_K y$ if and only if $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$.
- $x \prec_K y$ if and only if $\lambda^T x < \lambda^T y$ for all $\lambda \succeq_{K^*} 0, \lambda \neq 0$.

$K = K^{**}$ 이고 $\preceq_K^*$ 와 연관된 dual generalized inequality는 $\preceq_K$ 이기 때문에, generalized inequality와 dual이 바뀌더라도 이런 속성은 유지된다.

예를 들어서, $\lambda \preceq_K^* \mu$ if and only if $\lambda^T x \leq \mu^T x$ for all $x \succeq_K 0$ 이다.

01

02

Scalarization

Optimal for scalar problem ->Pareto-optimal for vector optimization problem

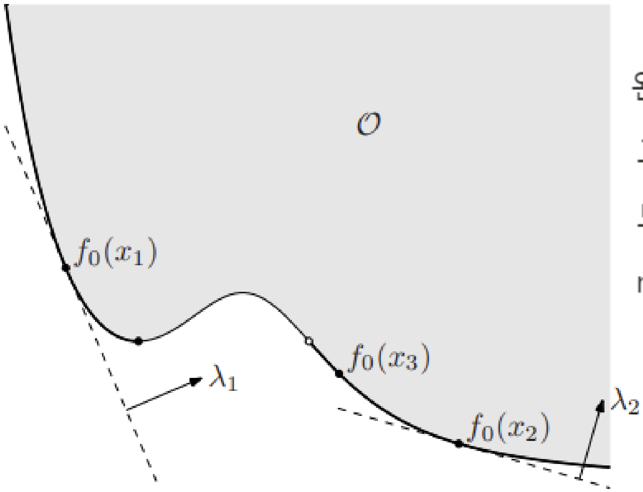
$\lambda \succ_{K^*} 0$

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

pf) If not, $f_0(y) \preceq_K f_0(x)$인데 $f_0(x) \neq f_0(y)$인 feasible y가 존재

$f_0(x) - f_0(y) \succeq_K 0$ 이고 nonzero이므로,

$\lambda^T f_0(x) > \lambda^T f_0(y)$인데 그러면 x가 optimal임에 모순



기하학적으로는 supporting hyperplane을 찾는 것

왼쪽 그림은 $f_0(x_1)$, $f_0(x_2)$는 scalarization으로 찾을 수 있으나, $f_0(x_3)$는 못 찾음

그러나, convex vector optimization problems에서는

모든 Pareto optimal point x^{po}에 대해,

nonzero $\lambda \succeq_{K^*} 0$가 존재하여 모든 Pareto optimal point를 찾을 수 있다

01

02

Multicriterion optimization

cone $K = R^q_+$ 인 vector optimization problem

$f_0(x) = (F_1(x), ..., F_q(x))$

Optimal

x^*가 optimal point면

$F_i(x^*) \geq F_i(y), i = 1, ..., q$를 모두 만족

Pareto Optimal

x^{po}가 Pareto optimal하면

feasible y에 대해 $F_i(y) \leq F_i(x)$이면 $F_i(y) = F_i(x^{po})$

Pareto optimal의 경우, trade-off가 일어남

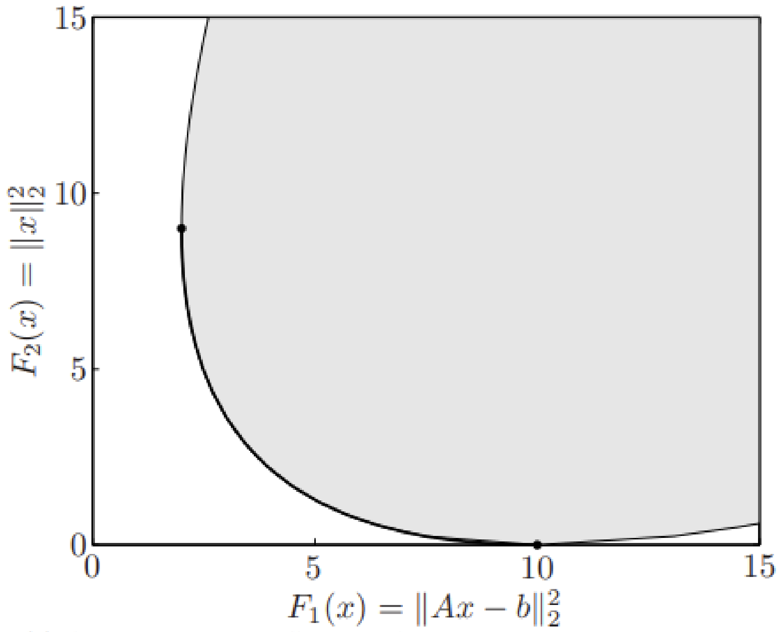
01

02

Scalarizing multicriterion problems

$\lambda^T f_0(x) = \lambda_1 F_1(x) + \cdots + \lambda_q F_q(x)$: weighted sum objective F_i를 줄이고 싶으면 λ_i를 크게 함

Regularized least-squares



take $\lambda = (1, \gamma), \gamma > 0$

minimize $\|Ax - b\|_2^2 + \gamma \|x\|_2^2$

과제

4.11 *Problems involving ℓ_1- and ℓ_∞-norms.* Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

(a) Minimize $\|Ax - b\|_\infty$ (ℓ_∞-norm approximation).

(b) Minimize $\|Ax - b\|_1$ (ℓ_1-norm approximation).

(c) Minimize $\|Ax - b\|_1$ subject to $\|x\|_\infty \leq 1$.

(d) Minimize $\|x\|_1$ subject to $\|Ax - b\|_\infty \leq 1$.

(e) Minimize $\|Ax - b\|_1 + \|x\|_\infty$.

In each problem, $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$ are given. (See §6.1 for more problems involving approximation and constrained approximation.)

감사합니다