Convex Optimization Problems

Yuanming Shi

ShanghaiTech University

Outline

Optimization Problems in Standard Form I

$$
\begin{array}{lll}
\text{min.} & \underset{\boldsymbol{x}}{\text{minimize}} & f_0(\boldsymbol{x}) \\
\text{s.t.} & \text{subject to} & f_i(\boldsymbol{x}) \le 0 \quad i = 1, \cdots, m \\
& & h_i(\boldsymbol{x}) = 0 \quad i = 1, \cdots, p
\end{array}
$$

minimize $f(x)$
subject to $x \in C$
feasible set

$C = \{x \mid f_i \le 0, h_i = 0\}$

minimize $f(x)$
$x \in C$

- $\boldsymbol{x} = (x_1, \cdots, x_n)$ is the optimization variable
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function
- $f_i : \mathbb{R}^n \to \mathbb{R} \quad i = 1, \cdots, m$ are the inequality constraint functions
- $h_i : \mathbb{R}^n \to \mathbb{R} \quad i = 1, \cdots, p$ are the equality constraint functions

Optimization Problems in Standard Form II

Feasibility:

- a point $\boldsymbol{x} \in \operatorname{dom} f_0$ is feasible if it satisfies all the constraints and infeasible otherwise
- a problem is feasible if it has at least one feasible point and infeasible otherwise.

Optimal value:

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

- $p^\star = +\infty$ if problem is infeasible (no $\boldsymbol{x}$ satisfies the constraints)
- $p^\star = -\infty$ if problem is unbounded below

Optimal solution: $\boldsymbol{x}^\star$ such that $f(\boldsymbol{x}^\star) = p^\star$ (and $x^\star$ feasible).

Global and Local Optimality

$X_{opt} = \{x \mid f_0(x) = p^\star\}$

- A feasible $\boldsymbol{x}$ is **optimal** if $f_0(\boldsymbol{x}) = p^\star$; X_{opt} is the set of optimal points.
- A feasible $\boldsymbol{x}$ is **locally optimal** if it is optimal within a ball, i.e., there is an $R > 0$ such that $\boldsymbol{x}^*$ is optimal for

$$\begin{array}{ll} \underset{\boldsymbol{z}}{\text{minimize}} & f_0(\boldsymbol{z}) \\ \text{subject to} & f_i(\boldsymbol{z}) \leq 0 \quad i = 1, \cdots, m \\ & h_i(\boldsymbol{z}) = 0 \quad i = 1, \cdots, p \\ & \|\boldsymbol{z} - \boldsymbol{x}^*\|_2 \leq R \end{array}$$

Example:

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

$\nabla f_0(x) = 0 \Rightarrow x = 1$

Implicit Constraints

- The standard form optimization problem has an explicit constraint:

$$\boldsymbol{x} \in \mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i \ \cap \ \bigcap_{i=1}^{p} \operatorname{dom} h_i$$

- $\mathcal{D}$ is the domain of the problem
- The constraints $f_i(\boldsymbol{x}) \leq 0, h_i(\boldsymbol{x}) = 0$ are the explicit constraints
- A problem is unconstrained if it has no explicit constraints
- Example:

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \log(b - \boldsymbol{a}^T\boldsymbol{x})$$

is an unconstrained problem with implicit constraint $b > \boldsymbol{a}^T\boldsymbol{x}$

Feasibility Problem

- Sometimes, we don't really want to minimize any objective, just to find a feasible point:

$$
\begin{array}{ll}
\underset{\boldsymbol{x}}{\text{find}} & \boldsymbol{x} \\
\text{subject to} & f_i(\boldsymbol{x}) \le 0 \quad i = 1, \cdots, m \\
 & h_i(\boldsymbol{x}) = 0 \quad i = 1, \cdots, p
\end{array}
$$

- This feasibility problem can be considered as a special case of a general problem:

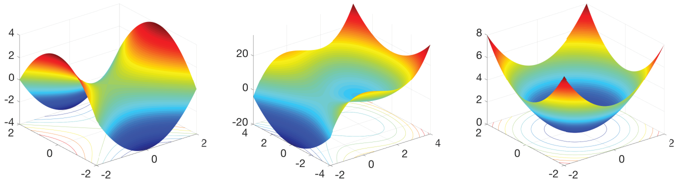
$$
\begin{array}{ll}
\underset{\boldsymbol{x}}{\text{minimize}} & 0 \quad c \\
\text{subject to} & f_i(\boldsymbol{x}) \le 0 \quad i = 1, \cdots, m \\
 & h_i(\boldsymbol{x}) = 0 \quad i = 1, \cdots, p
\end{array}
$$

where $p^\star = \underset{c}{0}$ if constraints are feasible and $p^\star \overset{+}{=} \infty$ otherwise.

Stationary Points

Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, a point $\boldsymbol{x} \in \mathbb{R}^n$ is called

- A **stationary point**, if $\nabla f(\boldsymbol{x}) = \mathbf{0}$;
- A **local minimum**, if $\boldsymbol{x}$ is a stationary point and there exists a neighborhood $\mathcal{B} \subseteq \mathbb{R}^n$ of $\boldsymbol{x}$ such that $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for any $\boldsymbol{y} \in \mathcal{B}$;
- A **global minimum**, if $\boldsymbol{x}$ is a stationary point and $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for any $\boldsymbol{y} \in \mathbb{R}^n$;
- **Saddle point**, if $\boldsymbol{x}$ is a stationary point and for any neighborhood $\mathcal{B} \subseteq \mathbb{R}^n$ of $\boldsymbol{x}$, there exist $\boldsymbol{y}, \boldsymbol{z} \in \mathcal{B}$ such that $f(\boldsymbol{z}) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ and $\lambda_{\min}(\nabla^2 f(\boldsymbol{x})) \leq 0$.



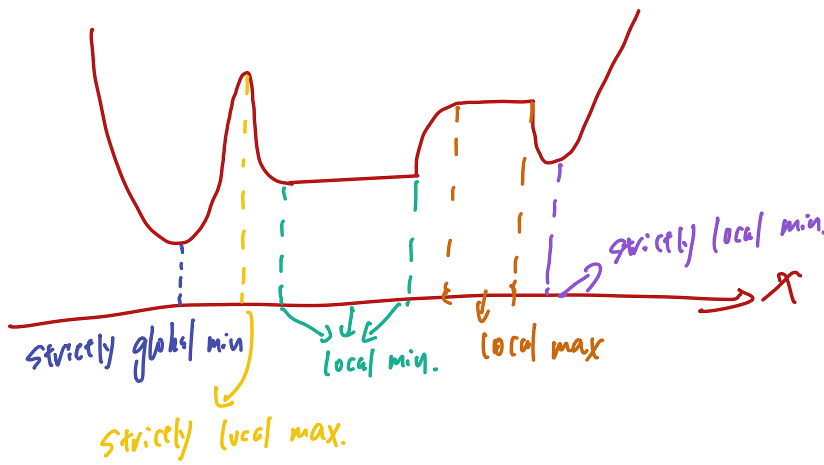
(a) strict saddle (b) local minimum (c) global minimum

$\lambda_{min} < 0$

Stationary points could be

$\begin{cases} \text{local minima, local maxima,} \\ \text{saddle points, global minima} \end{cases}$

A $\quad \lambda_{min}(\nabla^2 f(x)) \begin{cases} >0, & \text{local minimum } ✓ \\ =0, & \text{local minimum/ saddle points} \\ <0, & \text{strict saddle point } ✓ \end{cases}$



$$\begin{aligned} &\text{minimize} && f(x) \quad \text{—— (1)} \\ &\text{subject to} && x \in \Omega \rightarrow \text{feasible set} \end{aligned}$$

A point $\bar{x}$ is said to be a

① local minimum, if $\bar{x} \in \Omega$ and if $\exists\, \epsilon > 0$, s.t.

$f(\bar{x}) \leq f(x), \quad \forall\, x \in B(\bar{x}, \epsilon) \cap \Omega$

② strictly local minimum, if $\bar{x} \in \Omega$ and if $\bar{x} \in \Omega$, s.t.

$$f(\bar{x}) < f(x), \quad \forall x \in B(\bar{x}, \epsilon) \cap \Omega.$$

$$x \neq \bar{x}$$

③ global minimum if $\bar{x} \in \Omega$ and if

$$f(\bar{x}) \leq f(x), \quad \forall x \in \Omega$$

④ strictly global minimum if $\bar{x} \in \Omega$ and if

$$f(\bar{x}) < f(x), \quad \forall x \in \Omega, \ x \neq \bar{x}$$

Definition: A twice differentiable function $f(x)$ is strict saddle, if each critical point x^* $(\nabla f(x^*) = 0)$ of f is either

① a local minimizer with $\nabla^2 f(x) \succ 0$

② a "strict saddle" point, i.e., $\lambda_{\min}(\nabla^2 f(x)) < 0$

Q1: prove that "All local minimum are global optimal" (with a high probability)

1) if f is convex : easy!

2) if f is nonconvex : high-dimensional statistics optimization problem

① "matrix completion has no spurious local minima", by Rong Ge NIPS 2016, (student best paper)

$$\underset{M \in PSD}{\text{minimize}} \quad \sum_{(i,j) \in \Omega} (X_{ij} - M_{ij})^2$$

$$\text{subject to} \quad \operatorname{rank}(M) = r$$

$\Downarrow$

$$\underset{Y \in \mathbb{R}^{n \times r}}{\text{minimize}} \quad \sum_{(i,j) \in \Omega} \left(X_{ij} - (YY^T)_{ij}\right)^2$$

Q2 : All the saddle points can be escaped?

(in polynomial time complexity)

① "Gradien descent only converges to minimizers"

by Michael I. Jordan

Outline

First-order optimality conditions $\quad \underset{x}{\text{minimize}} \ f(x)$

Definition (Descent direction) Consider a function $f: \mathbb{R}^n \to \mathbb{R}$ and a point $x \in \mathbb{R}^n$. A direction $d \in \mathbb{R}^n$ is a descent direction at x, if $\exists \bar{\alpha} > 0$, s.t.

$$f(x + \alpha d) < f(x), \quad \alpha \in (0, \bar{\alpha}) \quad ②$$

Lemma 1: Consider a point $x \in \mathbb{R}^n$ and a continuously differentiable function f. Then, any direction d that satisfies

$$\boxed{\langle \nabla f(x), d \rangle < 0} \quad A \quad ①$$

is a descent direction

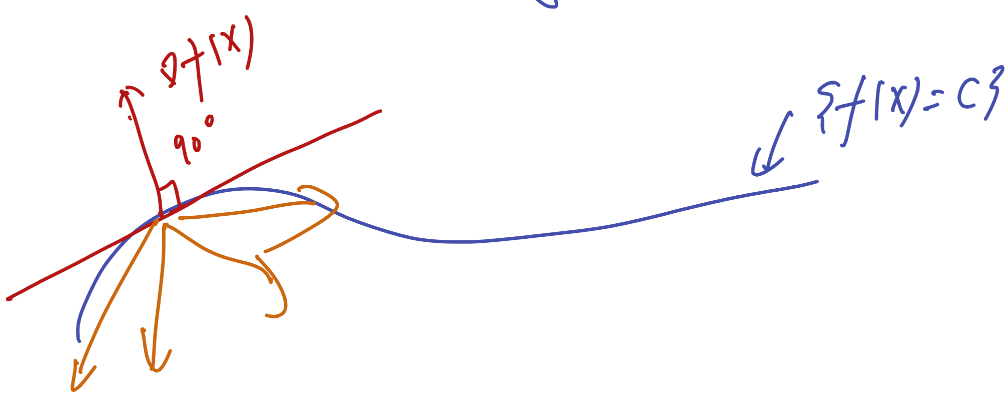
(In particular, $-\nabla f(x)$ is a descent direction if it is non-zero)

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

step size (learning rate in ML)

Gradient descent method

$$x_{k+1} = x_k - \alpha_k f(\nabla f(x_0), \ldots, \nabla f(x_k))$$



Proof: Let $g: \mathbb{R} \to \mathbb{R}$ be defined as

$g(\alpha) = f(x + \alpha d)$. Then

fixed fixed

$$g'(\alpha) = d^T \nabla f(x + \alpha d)$$

We use Taylor expansion to write

$$g(\alpha) = g(0) + g'(0)\alpha + o(\alpha)$$

$$\Longleftrightarrow f(x+\alpha d) = f(x) + \alpha d^T \nabla f(x) + \underline{o(\alpha)}$$

$$\Longleftrightarrow \frac{f(x+\alpha d) - f(x)}{\alpha} = \underbrace{d^T \nabla f(x)}_{<0} + \frac{o(\alpha)}{\alpha}$$

since $\lim\limits_{\alpha \downarrow 0} \frac{|o(\alpha)|}{\alpha} = 0$, there exists $\bar{\alpha}$, s.t.

$\forall \alpha \in (0, \bar{\alpha})$, we have $\frac{|o(\alpha)|}{\alpha} < \underbrace{\frac{1}{2}|d^T \nabla f(x)|}_{>0}$

Since $d^T \nabla f(x) < 0$ by assumption, we conclude that $\forall \alpha \in (0, \bar{\alpha})$,

$$f(x+\alpha d) - f(x) < \underline{\frac{1}{2} d^T \nabla f(x)} < 0$$

Remark: The converse of this lemma is not true.

e.g., $f(x) = x_1^2 - x_2^2$, $d = (0,1)^T$, $\bar{x} = (1,0)^T$,

for $\alpha \in \mathbb{R}$, we have

$f(\bar{x}+\alpha d) - f(\bar{x}) = 1^2 - (0+\alpha)^2 - 1^2 + 0 = -\alpha^2 < 0$

$\therefore d$ is a descent direction. but

$$\langle \nabla f(\bar{x}), d \rangle = (2,0)\cdot(0,1)^T = 0$$

First-order necessary condition for optimality

Theorem: If $\bar{x}$ is an unconstrained local minimum of a differentiable function $f: \mathbb{R}^n \to \mathbb{R}$, then

$$\nabla f(\bar{x}) = 0$$

Proof: if $\nabla f(\bar{x}) \neq 0$, then $\exists i$, s.t. $\frac{\partial f(x)}{\partial x_i} \neq 0$

then, from lemma 1, either e_i or $-e_i$ is a descent direction.

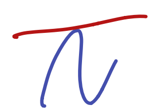
(Here, e_i is the i-th standard basis vector)

$$e_i = (0, \dots, 0, \underbrace{1}_{i\text{-th}}, 0, \dots, 0)$$

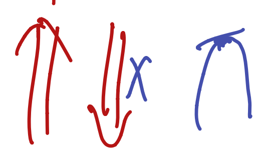
$$\frac{\partial f(x)}{\partial x_i} \begin{cases} > 0, & d = -e_i \Rightarrow \langle \nabla f(x), d \rangle = -\frac{\partial f(x)}{\partial x_i} < 0 \\ < 0, & d = e_i \Rightarrow \langle \nabla f(x), d \rangle = \frac{\partial f(x)}{\partial x_i} < 0 \end{cases}$$

Therefore, $\bar{x}$ cannot be a local minimum.

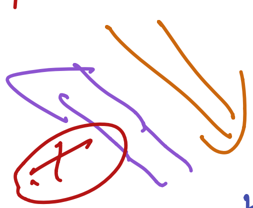
So far, we have



$$\bar{x} \text{ local minimum} \;\; \underset{\not\Leftarrow}{\Rightarrow} \;\; \nabla f(\bar{x}) = 0$$



example 1



$\nexists$ a descent direction at $\bar{x}$

example 1: Consider the function:

$$f(y,z) = (y^2 - z)(2y^2 - z)$$

$\boxed{\text{claim 1: } \underline{(0,0)} \text{ is a local minimum along every line that passes through it: } \nexists \text{ a descent direction at } (0,0)}$

Proof: For any direction $d = (d_1, d_2)^T$, let's look at

$$g(t) = f(td):$$

$$\underline{g(t)} = 2d_1^4 t^4 - 3d_1^2 d_2 t^3 + d_2^2 t^2$$

$$g'(t) = 8d_1^4 t^3 - 9d_1^2 d_2 t^2 + 2d_2^2 t$$

$$g''(t) = 24 d_1^4 t^2 - 18 d_1^2 d_2 t + 2d_2^2$$

$$g'(0) = 0 \quad , \quad g''(0) = 2d_2^2$$

Note that $g'(0) = 0$.

① if $d_2 \neq 0$, $t = 0$ is a strictly local minimum.

② if $d_2 = 0$, then $g(t) = \underline{2 d_1^4 t^4}$,

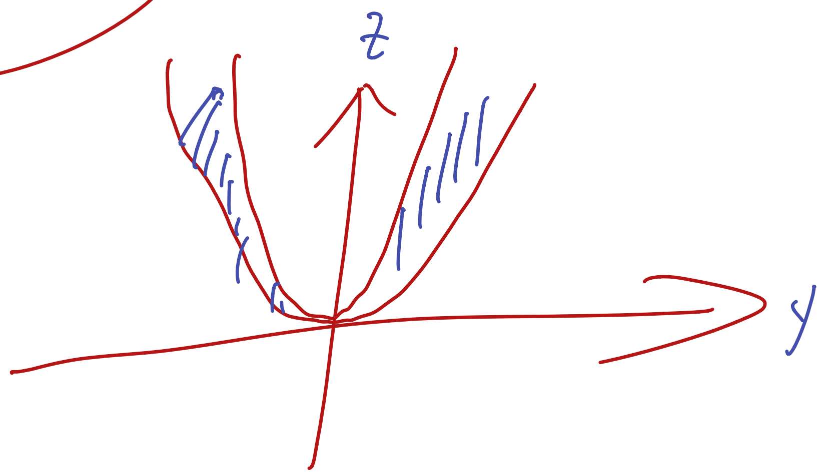
$t = 0$ is a strictly local minimum

Claim 2: $(0, 0)$ is not a local minimum

Proof: whenever $z \in (y^2, 2y^2)$,

the function $f(y, z) = (y^2 - z)(2y^2 - z)$ is

negative



$\boxed{\text{Second-order necessary condition for local optimality}}$

Theorem: If x^* is an unconstrained local minimizer of a twice differentiable function $f: \mathbb{R}^n \to \mathbb{R}$. Then, in addition to $\nabla f(x^*) = 0$, we must have

$$\lambda_{min}(\nabla^2 f(x^*)) \geq 0, \quad \nabla^2 f(x^*) \succeq 0$$

(positive semidefinite)

Proof: Consider some $y \in \mathbb{R}^n$. For $\alpha > 0$, the second-order Taylor expansion of f around x^* gives

$$f(x^* + \alpha y) = f(x^*) + \underbrace{\alpha y^T \nabla f(x^*)}_{=0} + \frac{\alpha^2}{2} y^T \nabla^2 f(x^*) y + o(\alpha^2)$$

$$\underbrace{\frac{f(x^*+\alpha y) - f(x^*)}{\alpha^2}}_{\geq 0} = \frac{1}{2} y^T \nabla^2 f(x^*) y + \frac{o(\alpha^2)}{\alpha^2}$$

By definition of local optimality of x^*, the left hand side is non-negative for α sufficiently small.

$$\Downarrow$$

$$\lim_{\alpha \downarrow 0} \frac{1}{2} y^T \nabla^2 f(x^*) y + \frac{o(\alpha^2)}{\alpha^2} \geq 0$$

But,

$$\lim_{\alpha \downarrow 0} \frac{o(\alpha^2)}{\alpha^2} = 0 \Rightarrow \underbrace{y}_{\text{can be arbitrary.}}{}^T \nabla^2 f(x^*) y \geq 0$$

$$\Downarrow$$

$$\nabla^2 f(x^*) \geq 0$$

Remark: The converse of this theorem is not true.

$\boxed{\text{Second-order sufficient condition for optimality}}$

Theorem: Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is twice differentiable and there exists a point of x^* such that

$$\boxed{Df(x^*) = 0, \quad \lambda_{min}(D^2 f(x^*)) > 0}$$

Then, x^* is strictly local minimum of f.

Proof: Let $\lambda > 0$ be the minimum eigenvalue of $D^2 f(x^*)$. This implies that

$$\underbrace{D^2 f(x^*) - \lambda I \succeq 0}$$

$$\underbrace{\lambda_1 - \lambda}_{\geq 0}, \ \cdots, \ \underbrace{\lambda_n - \lambda}_{= 0}$$

$$\underbrace{y^T D^2 f(x^*) y \geq \lambda \|y\|^2}, \quad \forall y \in \mathbb{R}^n$$

Once again, Taylor expansion yields

$$f(x^*+y) - f(x^*) = \underbrace{y^T \nabla f(x^*)}_{=0} +$$

$$\underbrace{\frac{1}{2} y^T \nabla^2 f(x^*) y}+ o(\|y\|^2)$$

$$\geq \frac{1}{2}\lambda \|y\|^2 + o(\|y\|^2)$$

$$= \|y\|^2 \left(\frac{\lambda}{2} + \frac{o(\|y\|^2)}{\|y\|^2} \right) \quad (y \neq 0)$$

Since $\lim_{\|y\| \to 0} \frac{o(\|y\|^2)}{\|y\|^2} = 0$, $\exists\, \delta > 0$, s.t.

$$\left| \frac{o(\|y\|^2)}{\|y\|^2} \right| < \frac{\lambda}{2}, \quad \forall y \text{ with } \|y\| < \delta.$$

Therefore, $f(x^*+y) > f(x^*)$, $\forall \|y\| < \delta$

x^* is a strict local minimum

Remark: the converse of this theorem is not true

$$\boxed{\text{Escape from the saddle points}}$$

$$\boxed{\lambda_{min}(\nabla^2 f(x)) < 0}$$

If we are at a strict saddle point, we need to consider a second-order Taylor expansion:

$$f(x + \Delta x) = f(x) + \underbrace{\Delta x^T \nabla f(x)}_{=0} + \frac{1}{2} \Delta x^T \nabla^2 f(x) \Delta x + o(\|\Delta x\|^2)$$

$$\boxed{\begin{array}{l}\text{We choose a descent direction} \\ \nabla^2 f(x) \Delta x = \lambda_{min}(\nabla^2 f(x)) \Delta x\end{array}}$$

$$\leq \frac{1}{2} \lambda_{min} \|\Delta x\|^2 + o(\|\Delta x\|^2)$$

$$\lim_{\Delta x \downarrow 0} \frac{o(\|\Delta x\|^2)}{\|\Delta x\|^2} = 0 \Rightarrow \left| \frac{o(\|\Delta x\|^2)}{\|\Delta x\|^2} \right| < \frac{-\lambda_{min}}{4}$$

$$f(x + \Delta x) - f(x) \leq \frac{1}{4} \lambda_{min} \|\Delta x\|^2$$

$$< 0$$

$\therefore$ we make local improvement along $\underline{\Delta x}$ $\nearrow \lambda_{min}$

as long as we have access to second-order information

Convex Optimization Problem

- Convex optimization problem in standard form:

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

  where $f_0, f_1, \cdots, f_m$ are convex and equality constraints are affine.

- **Local and global optima:** any locally optimal point of a convex problem is globally optimal

- Most problems are not convex when formulated

- Reformulating a problem in convex form is an art, there is no systematic way

Example

- The following problem is nonconvex (why not?):

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & x_1^2 + x_2^2 \\ \text{subject to} & x_1/(1+x_2^2) \leq 0 \\ & (x_1+x_2)^2 = 0 \end{array}$$

- The objective is convex.
- The equality constraint function is not affine; however, we can rewrite it as $x_1 = -x_2$ which is then a linear equality constraint.
- The inequality constraint function is not convex; however, we can rewrite it as $x_1 \leq 0$ which again is linear.
- We can rewrite it as

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 = -x_2 \end{array}$$

Global and Local Optimality

Any locally optimal point of a convex problem is globally optimal.
Proof: Suppose $\boldsymbol{x}$ is locally optimal (around a ball of radius R) and $\boldsymbol{y}$ is optimal with $f_0(\boldsymbol{y}) < f_0(\boldsymbol{x})$. We will show this cannot be.

Just take the segment from $\boldsymbol{x}$ to $\boldsymbol{y}$: $\boldsymbol{z} = \theta \boldsymbol{y} + (1-\theta)\boldsymbol{x}$.
Obviously the objective function is strictly decreasing along the segment since $f_0(\boldsymbol{y}) < f_0(\boldsymbol{x})$:

$$\theta f_0(\boldsymbol{y}) + (1-\theta) f_0(\boldsymbol{x}) < f_0(\boldsymbol{x}) \qquad \theta \in (0,1]$$

Using now the convexity of the function, we can write

$$f_0(\theta \boldsymbol{y} + (1-\theta)\boldsymbol{x}) < f_0(\boldsymbol{x}) \qquad \theta \in (0,1]$$

Finally, just choose θ sufficiently small such that the point $\boldsymbol{z}$ is in the ball of local optimality of $\boldsymbol{x}$, arriving at a contradiction.

convex optimization :

minimize $\quad f(x)$

subject to $\quad x \in \Omega$

f: convex function , Ω is a convex set

Theorem: Any local minimum is also a globally minimum.

Proof: Let $\bar{x}$ be a local minimum $\Rightarrow$

$\bar{x} \in \Omega$, and $\exists\, \epsilon > 0$, s.t.

$f(\bar{x}) \le f(x), \quad \forall x \in B(\bar{x}, \epsilon)$

Suppose for the sake of contradiction that $\exists\, z \in \Omega$ with

$$f(z) < f(\bar{x}) \qquad (*) \;—\; 1$$

Because of convexity of Ω, we have

$\lambda \bar{x} + (1-\lambda) z \in \Omega$, $\quad \forall \lambda \in (0,1)$

Because of convexity of f, we have

$$f(\lambda \bar{x} + (1-\lambda) z) \overset{\text{def.}}{\leq} \lambda f(\bar{x}) + (1-\lambda) f(z)$$

$$\overset{①}{<} \lambda f(\bar{x}) + (1-\lambda) f(\bar{x})$$

$$= f(\bar{x})$$

①

as $\lambda \to 1$, $\lambda \bar{x} + (1-\lambda) z \to B(\bar{x}, \epsilon)$

contradict local optimality of $\bar{x}$.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

Theorem: f is differentiable and convex. Any point $\bar{x}$ that satisfies $\nabla f(\bar{x}) = 0$ is a global minimum.

Proof: from the first-order characterization of convexity, we have

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y$$

In particular,

$$f(y) \geq f(\bar{x}) + \underbrace{\nabla f(\bar{x})}_{=0}{}^T (y - \bar{x}), \quad \forall y$$

$$\Downarrow$$

$$f(y) \geq f(\bar{x}), \quad \forall y.$$

Optimality Criterion for Differentiable f_0 I

Minimum Principle: A feasible point $\boldsymbol{x}$ is optimal if and only if

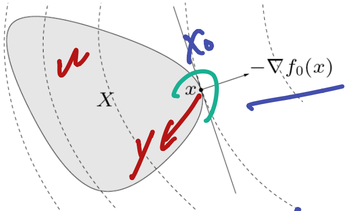
$$\nabla f_0(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}) \geq 0 \qquad \text{for all feasible } \boldsymbol{y} \in u$$

$\text{minimize}_{x \in \mathbb{R}^n} \; f_0(x)$

$\Rightarrow \nabla_0 f(x) = 0$

defines a hyperplane at x_0

$\{z \mid -\nabla f_0(x) z = -\nabla f_0(x) x_0\}$



$\downarrow \{x \mid f_0(x) = c\}$

$y = x - t\nabla f_0(x)$

$\Rightarrow -t\|\nabla f_0(x)\|^2 \geq 0 \Rightarrow \nabla f_0(x) = 0$

Optimality conditions for convex optimization.

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in \Omega \end{aligned}$$

$f: \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable and Ω is a convex set. Then a point is globally optimal iff $x \in \Omega$, and

$$\langle \nabla f(x), y - x \rangle \geq 0, \quad \forall y \in \Omega$$

Proof:

(sufficiency) Suppose $x \in \Omega$ satisfies

$$\langle \nabla f(x), y - x \rangle \geq 0, \quad \forall y \in \Omega \quad \text{— ①}$$

By the first-order characterization of convexity, we have

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall y \in \Omega \quad \text{— ②}$$

① + ② $\Rightarrow f(y) \geq f(x), \ \forall y \in \Omega$

$\Rightarrow x$ is globally optimal

(necessary) Suppose x is globally optimal,

but for some $y \in \Omega$, we have [descent direction]

$\langle \nabla f(x), y - x \rangle < 0$ ✗

Consider $g(\alpha) = f(x + \alpha(y - x))$.

Ω is convex $\Rightarrow \forall \alpha \in [0,1], \ x + \alpha(y-x) \in \Omega$

observe that $g'(\alpha) = (y-x)^T \nabla f(x + \alpha(y-x))$

$\Downarrow \ \alpha = 0$

$g'(0) = (y-x)^T \nabla f(x) < 0$

$\Downarrow$

$\exists \delta > 0$, s.t., $g(t) < g(0)$, $\forall t \in [0, \delta]$

$$\Downarrow$$

$f(x + t(y - x)) < f(x)$, $\forall t \in (0, \delta]$

Optimality Criterion for Differentiable f_0 II

- **Unconstrained problem:** $\boldsymbol{x}$ is optimal iff

$$\boldsymbol{x} \in \operatorname{dom} f_0, \qquad \nabla f_0(\boldsymbol{x}) = 0$$

(Lect. 4)

- **Equality constrained problem:** $\min_{\boldsymbol{x}} f_0(\boldsymbol{x}) \quad \text{s.t. } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$
$\boldsymbol{x}$ is optimal iff

KKT Condition

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

- **Minimization over nonnegative orthant:** $\min_{\boldsymbol{x}} f_0(\boldsymbol{x}) \quad \text{s.t. } \boldsymbol{x} \succeq \boldsymbol{0}$ $\boldsymbol{x}$ is optimal iff

$$\boldsymbol{x} \in \operatorname{dom} f_0, \qquad \boldsymbol{x} \succeq 0, \begin{cases} \nabla_i f_0(\boldsymbol{x}) \geq 0 & x_i = 0 \\ \nabla_i f_0(\boldsymbol{x}) = 0 & x_i > 0 \end{cases}$$

minimize $f(x)$ subject to $x \geq 0$.

optimality condition

$\langle \nabla f(x), y - x \rangle \geq 0, \quad x \geq 0, \quad y \geq 0$

$\nabla f(x)^T y - \nabla f(x)^T x \geq 0$

① $\nabla f(x)^T y = \sum_i \nabla f(x_i) y_i \quad (y_i \geq 0)$

is unbounded below, unless $\nabla f(x) \geq 0$

$(-\infty)$

$\Downarrow$

$-\nabla f(x)^T x \geq 0$

$\Downarrow \quad \nabla f(x) \geq 0, \quad x \geq 0$

$\nabla f(x)^T x = 0$

$\Downarrow$

Strict convexity and uniqueness of optimal solution

minimize $f(x)$

subject to $x \in \Omega$.

where $f: \mathbb{R}^n \to \mathbb{R}$ is strictly convex on Ω,

Ω is a convex set.

Then the optimal solution (assuming it exists) must be unique.

Proof: Suppose there are two optimal solutions $x, y \in \Omega$, and $x, y \in \mathbb{R}^n$. This means that

$$f(x) = f(y) \leq f(z), \quad \forall z \in \Omega \quad -① $$

Consider $z = \frac{x+y}{2}$. By convexity of Ω, we have $z \in \Omega$. By strict convexity, we have

$$f(z) = f\left(\frac{x+y}{2}\right) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = f(x)$$

this contradicts ①

Equivalent Reformulations I

- **Eliminating/introducing equality constraints:**

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

is equivalent to

$$\begin{array}{ll} \underset{\boldsymbol{z}}{\text{minimize}} & f_0(\boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0) \\ \text{subject to} & f_i(\boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0) \leq 0 \quad i = 1, \cdots, m \end{array}$$

where $\boldsymbol{F}$ and $\boldsymbol{x}_0$ are such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow \boldsymbol{x} = \boldsymbol{F}\boldsymbol{z} + \boldsymbol{x}_0$ for some $\boldsymbol{z}$.

Equivalent Reformulations II

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

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & f_0(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{a}_i^T \boldsymbol{x} \leq b_i \quad i = 1, \cdots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{\boldsymbol{x}, \boldsymbol{s}}{\text{minimize}} & f_0(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{a}_i^T \boldsymbol{x} + s_i = b_i \quad i = 1, \cdots, m \\ & s_i \geq 0 \end{array}$$

Equivalent Reformulations III

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

$$
\begin{array}{ll}
\underset{\boldsymbol{x},t}{\text{minimize}} & t \\
\text{subject to} & \underline{f_0(\boldsymbol{x}) - t \le 0} \\
 & f_i(\boldsymbol{x}) \le 0 \quad i = 1, \cdots, m \\
 & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{array}
$$

convex set

minimize $f_0(x)$

subject to $f_i(x) \le 0, \; i=1, \cdots, m$

$Ax = b$

Equivalent Reformulations IV

- **Minimizing over some variables:**

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & \underline{f_0(x,y)} \\ \text{subject to} & f_i(x) \leq 0 \quad i = 1, \cdots, m \end{array}$$

convex in (x,y)

is equivalent to

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \tilde{f}_0(x) \\ \text{subject to} & f_i(x) \leq 0 \quad i = 1, \cdots, m \end{array}$$

where $\underbrace{\tilde{f}_0(x) = \inf_y f_0(x,y)}_{\text{convex}}$ convex in (x, y)

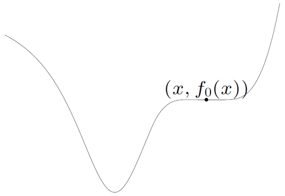
Outline

Quasiconvex Optimization

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

where $f_0 : \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasiconvex and $f_1, \cdots, f_m$ are convex

- Observe that it can have locally optimal points that are not (globally) optimal:



Quasiconvex Optimization

- **Convex representation** of sublevel sets of a quasiconvex function f_0: there exists a family of convex functions $\phi_t(\boldsymbol{x})$ for fixed t such that

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

- **Example:**

minimize
x

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

with p convex, q concave, and $p(\boldsymbol{x}) \ge 0$, $q(\boldsymbol{x}) > 0$ on $\mathrm{dom}\, f_0$. We can choose:

fixed

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

minimize$_x$ t
subject to $f_0(x) \le t$

  - for $t \ge 0$, $\phi_t(\boldsymbol{x})$ is convex in $\boldsymbol{x}$
  - $p(\boldsymbol{x})/q(\boldsymbol{x}) \le t$ if and only if $\phi_t(\boldsymbol{x}) \le 0$

convex

minimize $f_0(x)$ → quasi cvx.

subject to $\begin{cases} f_i(x) \le 0 \\ Ax = b \end{cases}$

cvx

Quasiconvex Optimization

Solving a quasiconvex problem via convex feasibility problems: the idea is to solve the epigraph form of the problem with a sandwich technique in t:

- for fixed t the epigraph form of the original problem reduces to a feasibility convex problem

$$\phi_t(\boldsymbol{x}) \le 0, \quad f_i(\boldsymbol{x}) \le 0 \forall i, \quad \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$$

- if t is too small, the feasibility problem will be infeasible
- if t is too large, the feasibility problem will be feasible
- start with upper and lower bounds on t (termed u and l, resp.) and use a sandwich technique (bisection method): at each iteration use $t = (l+u)/2$ and update the bounds according to the feasibility or infeasibility of the problem.

fixed t, fixed

minimize t

subject to $\begin{cases} f_0(x) \le t \\ f_i(x) \le 0 \\ Ax = 0 \end{cases}$

cvx.

cvx

feasible, $t_1 = \frac{l+t_0}{2}$

l t_0 u

$t_0 = \frac{l+u}{2}$

infeasible, $t_1 = \frac{t_0+u}{2}$

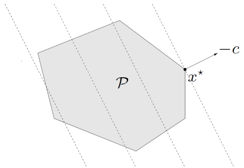
steps of bi-section: $\log_2 \left\lceil \frac{u-l}{\epsilon} \right\rceil$

Outline

Linear Programming (LP)

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

- Convex problem: affine objective and constraint functions.
- Feasible set is a polyhedron:



ℓ_1- and ℓ_∞- Norm Problems as LPs I

- ℓ_∞**-norm minimization:**

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{x}\|_\infty := \max_i |x_i| \\ \text{subject to} & \boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

$\Big\downarrow$ epigraph

is equivalent to the LP

$$\begin{array}{ll} \underset{t,\boldsymbol{x}}{\text{minimize}} & t \\ \text{subject to} & -t\mathbf{1} \preceq \boldsymbol{x} \preceq t\mathbf{1} \\ & \boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

$\|x\|_\infty \leq t$

$\Updownarrow$

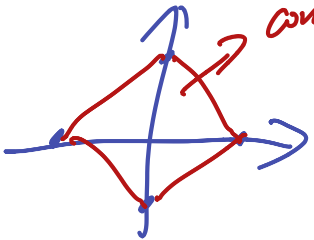
$\max_i |x_i| \leq t$

$\Updownarrow$

$|x_i| \leq t, \; i=1,\dots,n$

ℓ_1- and ℓ_∞- Norm Problems as LPs II

convex hull



- **ℓ_1-norm minimization:**

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{x}\|_1 \\ \text{subject to} & \boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

C

is equivalent to the LP

$$\min_{x \in C} \sum_{i=1}^{n} t_i$$

$$\text{s.t. } |x_i| = t_i, \; i = 1, \ldots, n$$

$$\begin{array}{ll} \underset{\boldsymbol{t},\boldsymbol{x}}{\text{minimize}} & \sum_i t_i \\ \text{subject to} & -\boldsymbol{t} \preceq \boldsymbol{x} \preceq \boldsymbol{t} \\ & \boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

(HW)

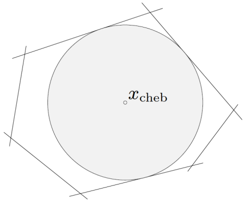
Examples: Chebyshev Center of a Polyhedron I

- Chebyshev center of a polyhedron

$$\mathcal{P} = \{\boldsymbol{x} \mid \boldsymbol{a}_i^T \boldsymbol{x} \le b_i,\ i = 1, \cdots, m\}$$

is center of largest inscribed ball

$$\mathcal{B} = \{\boldsymbol{x}_c + \boldsymbol{u} \mid \|\boldsymbol{u}\| \le r\}$$?



- Let's solve the problem

$$\begin{array}{lll} \underset{r, \boldsymbol{x}_c}{\text{maximize}} & r & \\ \text{subject to} & \boldsymbol{x} \in \mathcal{P} & \text{for all } \ \boldsymbol{x} = \boldsymbol{x}_c + \boldsymbol{u} \text{ with } \|\boldsymbol{u}\| \le r \end{array}$$

- Observe that $\boldsymbol{a}_i^T \boldsymbol{x} \le b_i$ for all $\boldsymbol{x} \in \mathcal{B}$ if and only if

$$\sup_{\boldsymbol{u}} \{\boldsymbol{a}_i^T \underbrace{(\boldsymbol{x}_c + \boldsymbol{u})}_{x} \mid \|\boldsymbol{u}\| \le r\} \le b_i$$

$$a_i^T x_c + a_i^T u \le \|a_i\| \, \|u\| \le r \|a_i\|$$

Examples: Chebyshev Center of a Polyhedron II

- Using Schwartz inequality, the supremum condition can be rewritten as

$$\boldsymbol{a}_i^T \boldsymbol{x}_c + r\|\boldsymbol{a}_i\|_2 \leq b_i$$

- Hence, the Chebyshev center can be obtained by solving:

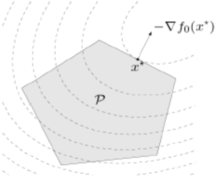
$$\begin{array}{ll} \underset{r, \boldsymbol{x}_c}{\text{maximize}} & r \\ \text{subject to} & \boldsymbol{a}_i^T \boldsymbol{x}_c + r\|\boldsymbol{a}_i\|_2 \leq b_i, \quad i = 1, \cdots, m \end{array}$$

which is an LP.

Quadratic Programming (QP)

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

- Convex problem (assuming $\boldsymbol{P} \in \mathbb{S}^n \succeq \mathbf{0}$): convex quadratic objective and affine constraint functions.
- Minimization of a convex quadratic function over a polyhedron:



Quadratically Constrained QP (QCQP)

$$
\begin{array}{lll}
\underset{\boldsymbol{x}}{\text{minimize}} & (1/2)\, \boldsymbol{x}^T \boldsymbol{P}_0 \boldsymbol{x} + \boldsymbol{q}_0^T \boldsymbol{x} + r_0 & \\
\text{subject to} & (1/2)\, \boldsymbol{x}^T \boldsymbol{P}_i \boldsymbol{x} + \boldsymbol{q}_i^T \boldsymbol{x} + r_i \leq 0 & i = 1, \cdots, m \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} &
\end{array}
$$

$= ?$

- Convex problem (assuming $\boldsymbol{P}_i \in \mathbb{S}^n \succeq \mathbf{0}$): convex quadratic objective and constraint functions.

$x^T P_i x = \mathrm{Trace}(P_i M)$

$M = x x^T, \quad \mathrm{rank}[M] = 1$

Second-Order Cone Programming (SOCP)

$$
\begin{array}{ll}
\underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{f}^T \boldsymbol{x} \\
\text{subject to} & \|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\| \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i \quad i = 1, \cdots, m \\
& \boldsymbol{F}\boldsymbol{x} = \boldsymbol{g}
\end{array}
$$

second-order cone

- Convex problem: linear objective and second-order cone constraints
- For $\boldsymbol{A}_i$ row vector, it reduces to an LP
- For $\boldsymbol{c}_i = \boldsymbol{0}$, it reduces to a QCQP
- More general than QCQP and LP

$$
\begin{cases}
z_i = A_i x + b_i \\
y_i = c_i^T x + d_i
\end{cases}
$$

$(z_i, y_i) \in K$

Robust LP as an SOCP

- Sometimes, the parameters of an optimization problem are imperfect
- Consider the robust LP:

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

where $\mathcal{E}_i = \{\bar{\boldsymbol{a}}_i + \boldsymbol{P}_i\boldsymbol{u} \mid \|\boldsymbol{u}\| \le 1\}$ uncertainty set

- It can be rewritten as the SOCP:

$a_i^T x + u \underbrace{P_i^T x}$

$\le \|u\| \, \|P_i^T x\|$

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \bar{\boldsymbol{a}}_i^T\boldsymbol{x} + \|\boldsymbol{P}_i^T\boldsymbol{x}\|_2 \le b_i \quad i = 1, \cdots, m \end{array}$$

$\sup_{a_i \in E_i} a_i^T x = \sup \{ \underbrace{(\bar{a}_i + P_i u)^T x} \mid \|u\| \le 1 \}$

Generalized Inequality Constraints

- Convex problem with generalized inequality constraints:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & f_0(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{f}_i(\boldsymbol{x}) \preceq_{K_i} \mathbf{0} \quad i = 1, \cdots, m \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

  where f_0 is convex and f_i are K_i-convex w.r.t. proper cone K_i
- It has the same properties as a standard convex problem
- **Conic form problem:** special case with affine objective and constraints:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{F}\boldsymbol{x} + \boldsymbol{g} \preceq_K \mathbf{0} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

P_1 conic optimization

$$\min_{X} \; C^T X \;, \quad \text{s.t. } AX=b,\; X \in K$$

tractable conic optimization

① The nonnegative orthant, $R_+^n \Rightarrow$ LP

② The second-order cone, $Q^n = \{(x,t) \in R^{n+1}, \|x\|_2 \le t\} \Rightarrow$ SOCP

③ The semidefinite cone, $S_+^n = \{X \mid X = X^T \succeq 0\} \Rightarrow$ SDP

P_0 $\min f_0(x)$

s.t. $f_i(x) \le 0$, $Ax=b$

$$P_0 \xrightarrow[\text{lect. 1~3}]{\text{CVX}} P_1 \xrightarrow[\text{SCS, ADMM}]{\text{SDPT3, SeDuMi}} \text{solution}$$

SDPT3, SeDuMi: interior-point

P_0: $\geq$ lect. 5

ADMM: KKT condition, lect. 4

Semidefinite Programming (SDP)

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} & \underbrace{x_1 \boldsymbol{F}_1 + x_2 \boldsymbol{F}_2 + \cdots + x_n \boldsymbol{F}_n}_{}^{A(x)} \preceq \boldsymbol{G} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array}$$

$\rightarrow G - A(x) \succeq 0$

- Inequality constraint is called linear matrix inequality (LMI)
- Convex problem: linear objective and linear matrix inequality (LMI) constraints
- Observe that multiple LMI constraints can always be written as a single one

SDP I

- **LP and equivalent SDP:**

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \preceq \boldsymbol{b} \end{array} \qquad \begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} & \operatorname{diag}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \preceq \boldsymbol{0} \end{array}$$

- **SOCP and equivalent SDP:**

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{f}^T \boldsymbol{x} \\ \text{subject to} & \|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\| \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i, \quad i = 1, \cdots, m \end{array}$$

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{f}^T \boldsymbol{x} \\ \text{subject to} & \begin{bmatrix} (\boldsymbol{c}_i^T \boldsymbol{x} + d_i)\boldsymbol{I} & \boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i \\ \boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i & \boldsymbol{c}_i^T \boldsymbol{x} + d_i \end{bmatrix} \succeq \boldsymbol{0}, \quad i = 1, \cdots, m \end{array}$$

SDP II

- **Eigenvalue minimization:**

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

where $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + x_1 \boldsymbol{A}_1 + \cdots + x_n \boldsymbol{A}_n$, is equivalent to SDP

$$\begin{array}{ll} \underset{\boldsymbol{x},t}{\text{minimize}} & t \\ \text{subject to} & \boldsymbol{A}(\boldsymbol{x}) \preceq t\boldsymbol{I} \end{array}$$

- It follows from

$$\underbrace{\lambda_{\max}(\boldsymbol{A}(\boldsymbol{x})) \le t}_{\text{epigraph}} \iff \underbrace{\boldsymbol{A}(\boldsymbol{x}) \preceq t\boldsymbol{I}}_{\lambda_i - t \le 0}$$

Reference

Chapter 4 of:

- Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

CVX