Convex Optimization

(EE227A: UC Berkeley)

Lecture 5
(Optimization problems)

05 Feb, 2013

Suvrit Sra

Organizational

- Homeworks due in class: 2/14/2013
- No late homeworks will be accepted
- Team up for projects into groups of 3–4 (max)
- Talk to me if special concerns
- We're using Piazza for Q/A — sign up!
- Bspace has the rest (course material, links, etc.)

Challenge

Consider the following functions on strictly positive variables:

$$
\begin{aligned}
h_1(x) &:= \frac{1}{x} \\
h_2(x,y) &:= \frac{1}{x} + \frac{1}{y} - \frac{1}{x+y} \\
h_3(x,y,z) &:= \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - \frac{1}{x+y} - \frac{1}{y+z} - \frac{1}{x+z} + \frac{1}{x+y+z}
\end{aligned}
$$

- ♡ Prove that h_1, h_2, h_3, and in general h_n are convex!
- ♡ Prove that in fact each $1/h_n$ is concave
- ♡ Generalize to where denom. replaced by $g(x)$, $g(x+y)$, $g(x+y+z)$, etc.

Challenge

Consider the following functions on strictly positive variables:

$$
\begin{aligned}
h_1(x) &:= \frac{1}{x} \\
h_2(x,y) &:= \frac{1}{x} + \frac{1}{y} - \frac{1}{x+y} \\
h_3(x,y,z) &:= \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - \frac{1}{x+y} - \frac{1}{y+z} - \frac{1}{x+z} + \frac{1}{x+y+z}
\end{aligned}
$$

♡ Prove that h_1, h_2, h_3, and in general h_n are convex!

♡ Prove that in fact each $1/h_n$ is concave

♡ Generalize to where denom. replaced by $g(x)$, $g(x+y)$, $g(x+y+z)$, etc.

$\nabla^2 h_n(x) \succeq 0$ is not recommended 🙂

Optimization problems

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f_1 \cdots \cap \mathrm{dom}\, f_m\}\,. \end{aligned}$$

Optimization problems

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \leq i \leq m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t. } & f_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}. \end{aligned}$$

Henceforth, we drop condition on domains for brevity.

Optimization problems

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t. } & f_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}. \end{aligned}$$

Henceforth, we drop condition on domains for brevity.

- If f_i are **differentiable** — smooth optimization
- If any of the f_i is **non-differentiable** — nonsmooth optimization
- If all f_i are **convex** — convex optimization
- If $m = 0$, i.e., only f_0 is there — **unconstrained** minimization

Convex optimization problems

Standard form

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

Convex optimization problems

Standard form

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

Some observations

- All f_i are convex

Convex optimization problems

Standard form

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

Some observations

- All f_i are convex
- Direction of inequality $f_i(x) \leq 0$ **crucial**

Convex optimization problems

Standard form

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

Some observations

- All f_i are convex
- Direction of inequality $f_i(x) \leq 0$ **crucial**
- The only equality constraints we allow are affine

Convex optimization problems

Standard form

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

Some observations

- All f_i are convex
- Direction of inequality $f_i(x) \leq 0$ **crucial**
- The only equality constraints we allow are affine
- This ensures, set of feasible solutions is also **convex**

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**
- By **convention**, we set $p^* = +\infty$ for infeasible problems

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**
- By **convention**, we set $p^* = +\infty$ for infeasible problems
- If $p^* = -\infty$, we say problem is **unbounded below**.

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf\{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**
- By **convention**, we set $p^* = +\infty$ for infeasible problems
- If $p^* = -\infty$, we say problem is **unbounded below**.
- Example, $\min x$ on $\mathbb{R}$, or $\min -\log x$ on $\mathbb{R}_{++}$

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**
- By **convention**, we set $p^* = +\infty$ for infeasible problems
- If $p^* = -\infty$, we say problem is **unbounded below**.
- Example, $\min x$ on $\mathbb{R}$, or $\min -\log x$ on $\mathbb{R}_{++}$
- Sometimes **minimum doesn't exist** (as $x \to \pm\infty$)

Convex optimization problems

Def. We denote by $\mathcal{X}$ the **feasible set**

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid f_i(x) \leq 0,\ 1 \leq i \leq m,\ Ax = b\}.$$

Def. We denote by p^* the **optimal value** of the problem.

$$p^* := \inf \{f_0(x) \mid x \in \mathcal{X}\}$$

- If $\mathcal{X}$ is empty, we say problem is **infeasible**
- By **convention**, we set $p^* = +\infty$ for infeasible problems
- If $p^* = -\infty$, we say problem is **unbounded below**.
- Example, $\min x$ on $\mathbb{R}$, or $\min -\log x$ on $\mathbb{R}_{++}$
- Sometimes **minimum doesn't exist** (as $x \to \pm\infty$)
- Say $f_0(x) = 0$, problem is called **convex feasibility**

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \le f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \le f(x)$ for **all** $x \in \mathcal{X}$.

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.
- If $y \notin \operatorname{dom} f$, then by definition $f(y) = +\infty$; nothing to prove

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.
- If $y \notin \operatorname{dom} f$, then by definition $f(y) = +\infty$; nothing to prove
- $\mathcal{X}$ is cvx., so we have $x_\theta = \theta y + (1-\theta)x^* \in \mathcal{X}$ for $\theta \in (0,1)$

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.
- If $y \notin \operatorname{dom} f$, then by definition $f(y) = +\infty$; nothing to prove
- $\mathcal{X}$ is cvx., so we have $x_\theta = \theta y + (1-\theta)x^* \in \mathcal{X}$ for $\theta \in (0,1)$
- Since f is cvx, and $x^*, y \in \operatorname{dom} f$, we have

$$f(x_\theta) - f(x^*) \leq \theta(f(y) - f(x^*)).$$

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.
- If $y \notin \operatorname{dom} f$, then by definition $f(y) = +\infty$; nothing to prove
- $\mathcal{X}$ is cvx., so we have $x_\theta = \theta y + (1-\theta)x^* \in \mathcal{X}$ for $\theta \in (0,1)$
- Since f is cvx, and $x^*, y \in \operatorname{dom} f$, we have

$$f(x_\theta) - f(x^*) \leq \theta(f(y) - f(x^*)).$$

- Since x^* is a local minimizer, for small enough $\theta > 0$, lhs ≥ 0.

Optimality

Def. A point $x^* \in \mathcal{X}$ is **locally optimal** if $f(x^*) \leq f(x)$ for all x in a **neighborhood** of x^*. **Global** if $f(x^*) \leq f(x)$ for **all** $x \in \mathcal{X}$.

Theorem For convex problems, locally optimal point also global.

- Let x^* be a local minimizer of $f(x)$ on $\mathcal{X}$
- Let $y \in \mathcal{X}$ be any other feasible point.
- We need to show that $f(y) \geq f(x^*) = p^*$.
- If $y \notin \operatorname{dom} f$, then by definition $f(y) = +\infty$; nothing to prove
- $\mathcal{X}$ is cvx., so we have $x_\theta = \theta y + (1-\theta)x^* \in \mathcal{X}$ for $\theta \in (0,1)$
- Since f is cvx, and $x^*, y \in \operatorname{dom} f$, we have

$$f(x_\theta) - f(x^*) \leq \theta(f(y) - f(x^*)).$$

- Since x^* is a local minimizer, for small enough $\theta > 0$, lhs ≥ 0.
- So rhs is also nonnegative, proving $f(y) \geq f(x^*)$ as desired.

Optimal set

The set of optimal solutions $\mathcal{X}^*$ may be empty

Example If $\mathcal{X} = \emptyset$, i.e., no feasible solutions, then $\mathcal{X}^* = \emptyset$

Optimal set

The set of optimal solutions $\mathcal{X}^*$ may be empty

Example If $\mathcal{X} = \emptyset$, i.e., no feasible solutions, then $\mathcal{X}^* = \emptyset$

Example When only $\inf$ and not $\min$, e.g., $\inf e^x$ as $x \to -\infty$ or $\inf 1/x$ as $x \to \infty$; so sometimes we have to worry about $\mathcal{X}^* = \emptyset$

Optimal set

The set of optimal solutions $\mathcal{X}^*$ may be empty

Example If $\mathcal{X} = \emptyset$, i.e., no feasible solutions, then $\mathcal{X}^* = \emptyset$

Example When only $\inf$ and not $\min$, e.g., $\inf e^x$ as $x \to -\infty$ or $\inf 1/x$ as $x \to \infty$; so sometimes we have to worry about $\mathcal{X}^* = \emptyset$

Exercise: Verify that $\mathcal{X}^*$ is always convex.

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$.
Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \quad \frac{f(x^* + td) - f(x^*)}{}$$

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \frac{f(x^* + td) - f(x^*)}{t}$$

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \lim_{t \downarrow 0} \frac{f(x^* + td) - f(x^*)}{t}$$

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \lim_{t \downarrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \frac{dg(0)}{dt}$$

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \lim_{t \downarrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \frac{dg(0)}{dt} = \langle \nabla f(x^*), d \rangle.$$

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \lim_{t \downarrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \frac{dg(0)}{dt} = \langle \nabla f(x^*), d \rangle.$$

Similarly, using $-d$ it follows that $\langle \nabla f(x^*), d \rangle \leq 0$, so $\langle \nabla f(x^*), d \rangle = 0$ **must hold**. Since d is arbitrary, $\nabla f(x^*) = 0$.

First-order optimality conditions

Theorem Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in an open set S containing x^*, a local minimum of f. Then, $\nabla f(x^*) = 0$.

Proof: Consider function $g(t) = f(x^* + td)$, where $d \in \mathbb{R}^n$; $t > 0$. Since x^* is a local min, for small enough t, $f(x^* + td) \geq f(x^*)$.

$$0 \leq \lim_{t \downarrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \frac{dg(0)}{dt} = \langle \nabla f(x^*), d \rangle.$$

Similarly, using $-d$ it follows that $\langle \nabla f(x^*), d \rangle \leq 0$, so $\langle \nabla f(x^*), d \rangle = 0$ **must hold**. Since d is arbitrary, $\nabla f(x^*) = 0$.

Exercise: Prove that if f is convex, then $\nabla f(x^*) = 0$ is actually **sufficient** for global optimality! For general f this is **not** true. (This property that makes convex optimization special!)

Optimality conditions – constrained

♠ For every $x, y \in \operatorname{dom} f$, we have $f(y) \geq f(x) + \langle \nabla f(x),\, y - x\rangle$.

Optimality conditions – constrained

♠ For every $x, y \in \operatorname{dom} f$, we have $f(y) \geq f(x) + \langle \nabla f(x),\, y - x \rangle$.

♠ Thus, x^* is optimal if and only if

$$\langle \nabla f(x^*),\, y - x^* \rangle \geq 0, \qquad \text{for all } \; y \in \mathcal{X}.$$

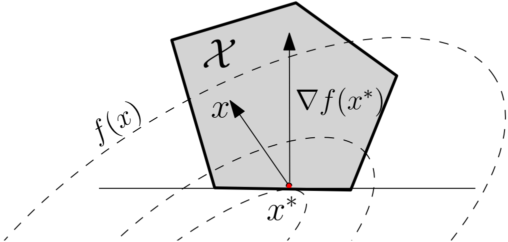
Optimality conditions – constrained

♠ For every $x, y \in \operatorname{dom} f$, we have $f(y) \geq f(x) + \langle \nabla f(x),\, y - x \rangle$.

♠ Thus, x^* is optimal if and only if

$$\langle \nabla f(x^*),\, y - x^* \rangle \geq 0, \qquad \text{for all } \; y \in \mathcal{X}.$$

♠ If $\mathcal{X} = \mathbb{R}^n$, this reduces to $\nabla f(x^*) = 0$



♠ If $\nabla f(x^*) \neq 0$, it defines supporting hyperplane to $\mathcal{X}$ at x^*

Optimality conditions – constrained

- Suppose $\exists y \in \mathcal{X}$ such that $\langle \nabla f(x^*), y - x^* \rangle < 0$
- Using mean-value theorem of calculus, $\exists \xi \in [0, 1]$ s.t.

$$f(x^* + t(y - x^*)) = f(x^*) + \langle \nabla f(x^* + \xi t(y - x^*)), t(y - x^*) \rangle$$

  (we applied MVT to $g(t) := f(x^* + t(y - x^*))$)
- For sufficiently small t, since ∇f continuous, by assump on y, $\langle \nabla f(x^* + \xi t(y - x^*)), y - x^* \rangle < 0$
- This in turn implies that $f(x^* + t(y - x^*)) < f(x^*)$
- Since $\mathcal{X}$ is convex, $x^* + t(y - x^*) \in \mathcal{X}$ is also feasible
- Contradiction to local optimality of x^*

Optimality – nonsmooth

Theorem (Fermat's rule): Let $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$. Then,

$$\operatorname{Argmin} f_0 = \mathsf{zer}(\partial f_0) := \{x \in \mathbb{R}^n \mid 0 \in \partial f_0(x)\}\,.$$

Optimality – nonsmooth

Theorem (Fermat's rule): Let $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$. Then,

$$\operatorname{Argmin} f_0 = \mathsf{zer}(\partial f_0) := \{x \in \mathbb{R}^n \mid 0 \in \partial f_0(x)\}.$$

Proof: $x \in \operatorname{Argmin} f_0$ implies that $f_0(x) \leq f_0(y)$ for all $y \in \mathbb{R}^n$.

Optimality – nonsmooth

Theorem (Fermat's rule): Let $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$. Then,

$$\operatorname{Argmin} f_0 = \mathsf{zer}(\partial f_0) := \{x \in \mathbb{R}^n \mid 0 \in \partial f_0(x)\}\,.$$

Proof: $x \in \operatorname{Argmin} f_0$ implies that $f_0(x) \leq f_0(y)$ for all $y \in \mathbb{R}^n$.
Equivalently, $f_0(y) \geq f_0(x) + \langle 0,\, y - x\rangle \quad \forall y$,

Optimality – nonsmooth

Theorem (Fermat's rule): Let $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$. Then,

$$\operatorname{Argmin} f_0 = \mathsf{zer}(\partial f_0) := \{x \in \mathbb{R}^n \mid 0 \in \partial f_0(x)\}.$$

Proof: $x \in \operatorname{Argmin} f_0$ implies that $f_0(x) \leq f_0(y)$ for all $y \in \mathbb{R}^n$.
Equivalently, $f_0(y) \geq f_0(x) + \langle 0,\, y - x\rangle \quad \forall y$, $\leftrightarrow 0 \in \partial f_0(x)$.

Optimality – nonsmooth

Theorem (Fermat's rule): Let $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$. Then,

$$\operatorname{Argmin} f_0 = \mathsf{zer}(\partial f_0) := \{x \in \mathbb{R}^n \mid 0 \in \partial f_0(x)\}.$$

Proof: $x \in \operatorname{Argmin} f_0$ implies that $f_0(x) \leq f_0(y)$ for all $y \in \mathbb{R}^n$.
Equivalently, $f_0(y) \geq f_0(x) + \langle 0,\, y - x\rangle \quad \forall y$, $\leftrightarrow 0 \in \partial f_0(x)$.

Nonsmooth optimality

$$\begin{aligned} \min \quad & f_0(x) \qquad \text{s.t. } x \in \mathcal{X} \\ \min \quad & f_0(x) + \mathbb{I}_{\mathcal{X}}(x). \end{aligned}$$

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$
- Recall, $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ iff $\mathbb{I}_{\mathcal{X}}(y) \geq \mathbb{I}_{\mathcal{X}}(x) + \langle g,\, y - x\rangle$ for all y.

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$
- Recall, $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ iff $\mathbb{I}_{\mathcal{X}}(y) \geq \mathbb{I}_{\mathcal{X}}(x) + \langle g,\, y - x \rangle$ for all y.
- So $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ means $x \in \mathcal{X}$ and $0 \geq \langle g,\, y - x \rangle$ $\forall y \in \mathcal{X}$.

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$
- Recall, $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ iff $\mathbb{I}_{\mathcal{X}}(y) \geq \mathbb{I}_{\mathcal{X}}(x) + \langle g,\, y - x\rangle$ for all y.
- So $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ means $x \in \mathcal{X}$ and $0 \geq \langle g,\, y - x\rangle \; \forall y \in \mathcal{X}$.
- **Normal cone:**

$$\mathcal{N}_{\mathcal{X}}(x) := \{g \in \mathbb{R}^n \mid 0 \geq \langle g,\, y - x\rangle \quad \forall y \in \mathcal{X}\}$$

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$
- Recall, $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ iff $\mathbb{I}_{\mathcal{X}}(y) \geq \mathbb{I}_{\mathcal{X}}(x) + \langle g,\, y - x\rangle$ for all y.
- So $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ means $x \in \mathcal{X}$ and $0 \geq \langle g,\, y - x\rangle$ $\forall y \in \mathcal{X}$.
- **Normal cone:**

$$\mathcal{N}_{\mathcal{X}}(x) := \{g \in \mathbb{R}^n \mid 0 \geq \langle g,\, y - x\rangle \quad \forall y \in \mathcal{X}\}$$

Application. $\min f_0(x)$ s.t. $x \in \mathcal{X}$:

♢ If f_0 is diff., we get $0 \in \nabla f(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$

Optimality – nonsmooth

- Minimizing x must satisfy: $0 \in \partial(f_0 + \mathbb{I}_{\mathcal{X}})(x)$
- (**CQ**) Assuming $\mathrm{ri}(\mathrm{dom}\, f_0) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, $0 \in \partial f_0(x) + \partial \mathbb{I}_X(x)$
- Recall, $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ iff $\mathbb{I}_{\mathcal{X}}(y) \geq \mathbb{I}_{\mathcal{X}}(x) + \langle g,\, y - x\rangle$ for all y.
- So $g \in \partial \mathbb{I}_{\mathcal{X}}(x)$ means $x \in \mathcal{X}$ and $0 \geq \langle g,\, y - x\rangle\ \forall y \in \mathcal{X}$.
- **Normal cone:**

$$\mathcal{N}_{\mathcal{X}}(x) := \{g \in \mathbb{R}^n \mid 0 \geq \langle g,\, y - x\rangle \quad \forall y \in \mathcal{X}\}$$

Application. $\min f_0(x)$ s.t. $x \in \mathcal{X}$:

♢ If f_0 is diff., we get $0 \in \nabla f(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$

♢ $-\nabla f(x^*) \in \mathcal{N}_{\mathcal{X}}(x^*) \iff \langle \nabla f(x^*),\, y - x^*\rangle \geq 0$ for all $y \in \mathcal{X}$.

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \le 1.}$$

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \leq 1.}$$

A point x is optimal if and only if

$x \in \operatorname{dom} f, \quad \|x\| \leq 1,$

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \leq 1.}$$

A point x is optimal if and only if

$$x \in \operatorname{dom} f, \quad \|x\| \leq 1, \forall y \text{ s.t. } \|y\| \leq 1 \implies \nabla f(x)^T (y - x) \geq 0.$$

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \leq 1.}$$

A point x is optimal if and only if

$$x \in \operatorname{dom} f, \quad \|x\| \leq 1, \forall y \text{ s.t. } \|y\| \leq 1 \implies \nabla f(x)^T(y - x) \geq 0.$$

In other words

$$\begin{aligned} \forall \|y\| \leq 1, \quad \nabla f(x)^T y &\geq \nabla f(x)^T x \\ \forall \|y\| \leq 1, \quad -\nabla f(x)^T y &\leq -\nabla f(x)^T x \end{aligned}$$

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \leq 1.}$$

A point x is optimal if and only if

$$x \in \operatorname{dom} f, \quad \|x\| \leq 1, \forall y \text{ s.t. } \|y\| \leq 1 \implies \nabla f(x)^T (y - x) \geq 0.$$

In other words

$$\begin{aligned} \forall \|y\| \leq 1, \quad \nabla f(x)^T y &\geq \nabla f(x)^T x \\ \forall \|y\| \leq 1, \quad -\nabla f(x)^T y &\leq -\nabla f(x)^T x \\ \sup\{-\nabla f(x)^T y \mid \|y\| \leq 1\} &\leq -\nabla f(x)^T x \end{aligned}$$

Example

$$\min \quad f(x) \qquad \|x\| \leq 1.$$

A point x is optimal if and only if

$$x \in \operatorname{dom} f, \quad \|x\| \leq 1, \forall y \text{ s.t. } \|y\| \leq 1 \implies \nabla f(x)^T(y - x) \geq 0.$$

In other words

$$\begin{aligned}
\forall \|y\| \leq 1, \quad \nabla f(x)^T y &\geq \nabla f(x)^T x \\
\forall \|y\| \leq 1, \quad -\nabla f(x)^T y &\leq -\nabla f(x)^T x \\
\sup\{-\nabla f(x)^T y \mid \|y\| \leq 1\} &\leq -\nabla f(x)^T x \\
\|-\nabla f(x)\|_* &\leq -\nabla f(x)^T x \\
\|\nabla f(x)\|_* &\leq -\nabla f(x)^T x.
\end{aligned}$$

Example

$$\boxed{\min \quad f(x) \qquad \|x\| \leq 1.}$$

A point x is optimal if and only if

$$x \in \operatorname{dom} f, \quad \|x\| \leq 1, \forall y \text{ s.t. } \|y\| \leq 1 \implies \nabla f(x)^T (y - x) \geq 0.$$

In other words

$$\begin{array}{rcl}
\forall \|y\| \leq 1, \quad \nabla f(x)^T y & \geq & \nabla f(x)^T x \\
\forall \|y\| \leq 1, \quad -\nabla f(x)^T y & \leq & -\nabla f(x)^T x \\
\sup\{-\nabla f(x)^T y \mid \|y\| \leq 1\} & \leq & -\nabla f(x)^T x \\
\|-\nabla f(x)\|_* & \leq & -\nabla f(x)^T x \\
\|\nabla f(x)\|_* & \leq & -\nabla f(x)^T x.
\end{array}$$

Observe: If constraint not satisfied strictly at optimum ($\|x\| < 1$), then $\nabla f(x) = 0$ (else we'd violate the last inequality above).

Equivalent Problems

Monotonic transformation

Standard form

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

Monotonic transformation

Standard form

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

- Say $\psi_0 : \mathbb{R} \to \mathbb{R}$ is monotone increasing
- $\psi_i : \mathbb{R} \to \mathbb{R}$ satisfy $\psi_i(u) \leq 0$ iff $u \leq 0$
- $h(z) = 0$ iff $z = 0$.

Monotonic transformation

Standard form

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

- Say $\psi_0 : \mathbb{R} \to \mathbb{R}$ is monotone increasing
- $\psi_i : \mathbb{R} \to \mathbb{R}$ satisfy $\psi_i(u) \le 0$ iff $u \le 0$
- $h(z) = 0$ iff $z = 0$.

Transformed problem

$$\begin{aligned} \min \quad & \psi_0(f_0(x)) \\ \text{s.t.} \quad & \psi_i(f_i(x)) \le 0, \quad i = 1, \ldots, m \\ & h(Ax - b) = 0. \end{aligned}$$

Monotonic transformation

Standard form

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

- Say $\psi_0 : \mathbb{R} \to \mathbb{R}$ is monotone increasing
- $\psi_i : \mathbb{R} \to \mathbb{R}$ satisfy $\psi_i(u) \leq 0$ iff $u \leq 0$
- $h(z) = 0$ iff $z = 0$.

Transformed problem

$$\begin{aligned} \min \quad & \psi_0(f_0(x)) \\ \text{s.t.} \quad & \psi_i(f_i(x)) \leq 0, \quad i = 1, \ldots, m \\ & h(Ax - b) = 0. \end{aligned}$$

$\boxed{\text{Can destroy convexity}}$

Example

$$
\begin{aligned}
\min \qquad & \|Ax - b\| \\
\min \qquad & \|Ax - b\|^2
\end{aligned}
$$

Example

$$\begin{aligned} &\min \quad \|Ax - b\| \\ &\min \quad \|Ax - b\|^2 \end{aligned}$$

- ♣ Set of optimal points same
- ♣ Problems equivalent but not same

Example

$$\begin{aligned} \min \quad & \|Ax - b\| \\ \min \quad & \|Ax - b\|^2 \end{aligned}$$

- ♣ Set of optimal points same
- ♣ Problems equivalent but not same
- ♣ First problem is nondifferentiable
- ♣ Second is differentiable – solvable in closed form!

Slack variables

To turn inequalities into equalities

$$\min \quad f(x) \quad \text{s.t.} \quad Ax \leq b$$

$$\min_{x,s} \quad f(x) \quad \text{s.t.} \quad Ax + s = b, \; s \geq 0.$$

Epigraph form

Standard form; optimal value p^*

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

Epigraph form

Standard form; optimal value p^*

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

Epigraph form

$$\begin{aligned} \min_{(x,t)} \quad & t \\ \text{s.t. } & f_0(x) - t \le 0 \\ & f_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

At the optimum, $t = p^*$.

Epigraph form

Standard form; optimal value p^*

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

Epigraph form

$$\begin{aligned} \min_{(x,t)} \quad & t \\ \text{s.t. } & f_0(x) - t \le 0 \\ & f_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

At the optimum, $t = p^*$.
In other words: Define **sublevel set** $L_t := \{x \mid f_0(x) \le t\}$, $t \in \mathbb{R}$.

Epigraph form

Standard form; optimal value p^*

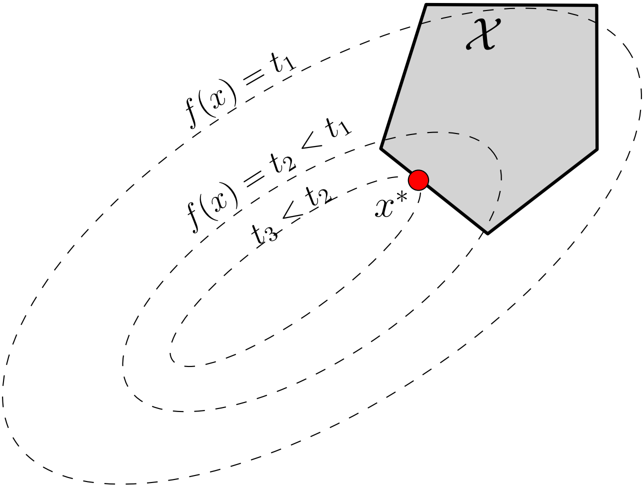
$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

Epigraph form

$$\begin{aligned} \min_{(x,t)} \quad & t \\ \text{s.t.} \quad & f_0(x) - t \le 0 \\ & f_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

At the optimum, $t = p^*$.
In other words: Define **sublevel set** $L_t := \{x \mid f_0(x) \le t\}$, $t \in \mathbb{R}$.
We seek smallest t such that L_t intersects with constraint set $\mathcal{X}$.

Epigraph form — geometrically



Variable elimination

$$\min_{x,y} \quad f_0(x,y) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1,\ldots,m.$$

Recall, since f_0 is convex in (x,y), $\inf_y f_0(x,y)$ is still convex.

Variable elimination

$$\min_{x,y} \quad f_0(x,y) \quad \text{s.t.} \;\; f_i(x) \le 0, \quad i = 1,\dots,m.$$

Recall, since f_0 is convex in (x,y), $\inf_y f_0(x,y)$ is still convex.

Variable elimination

$$\min_{x} \quad \tilde{f}_0(x) \quad \text{s.t.} \;\; f_i(x) \le 0, \quad i = 1,\dots,m$$

where $\tilde{f}_0(x) := \inf_y f_0(x,y)$.

Variable elimination

$$\min_{x,y} \quad f_0(x, y) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1, \ldots, m.$$

Recall, since f_0 is convex in (x, y), $\inf_y f_0(x, y)$ is still convex.

Variable elimination

$$\min_{x} \quad \tilde{f}_0(x) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1, \ldots, m$$

where $\tilde{f}_0(x) := \inf_y f_0(x, y)$.

More generally: $\tilde{f}_0(x) := \inf \{f_0(x, y) \mid g_i(y) \le 0, i = 1, \ldots, m'\}$

Variable elimination

$$\min_{x,y} \quad f_0(x,y) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1,\ldots,m.$$

Recall, since f_0 is convex in (x,y), $\inf_y f_0(x,y)$ is still convex.

Variable elimination

$$\min_{x} \quad \tilde{f}_0(x) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1,\ldots,m$$

where $\tilde{f}_0(x) := \inf_y f_0(x,y)$.

More generally: $\tilde{f}_0(x) := \inf \{ f_0(x,y) \mid g_i(y) \le 0, i = 1,\ldots,m' \}$

$\boxed{\text{Independent constraints important here.}}$

Equality constraint elimination

Standard form

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

Equality constraint elimination

Standard form

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

- For $Ax = b$ to be feasible, $b \in \mathcal{R}(A)$.

Equality constraint elimination

Standard form

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

- For $Ax = b$ to be feasible, $b \in \mathcal{R}(A)$.
- Let x_0 be any solution to $Ax = b$.

Equality constraint elimination

Standard form

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

- For $Ax = b$ to be feasible, $b \in \mathcal{R}(A)$.
- Let x_0 be any solution to $Ax = b$.
- Let F be a matrix with $\mathcal{R}(F) = \mathcal{N}(A)$; so $AFz = 0$

Equality constraint elimination

Standard form

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

- For $Ax = b$ to be feasible, $b \in \mathcal{R}(A)$.
- Let x_0 be any solution to $Ax = b$.
- Let F be a matrix with $\mathcal{R}(F) = \mathcal{N}(A)$; so $AFz = 0$
- General solution to $Ax = b$ is of form: $Fz + x_0$

Equality constraint elimination

Standard form

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

- For $Ax = b$ to be feasible, $b \in \mathcal{R}(A)$.
- Let x_0 be any solution to $Ax = b$.
- Let F be a matrix with $\mathcal{R}(F) = \mathcal{N}(A)$; so $AFz = 0$
- General solution to $Ax = b$ is of form: $Fz + x_0$

Elimination form

$$\begin{aligned} \min \quad & f_0(Fz + x_0) \\ \text{s.t.} \quad & f_i(Fz + x_0) \le 0, \quad 1 \le i \le m. \end{aligned}$$

Introducing equality constraints

Separable function

$$\begin{array}{ll} \min & \sum_{i=1}^{T} f_{0,i}(x) \\ \text{s.t.} & f_i(x) \le 0, \quad i = 1, \ldots, m. \end{array}$$

Introducing equality constraints

Separable function

$$\begin{aligned} \min \quad & \sum_{i=1}^{T} f_{0,i}(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad i = 1, \ldots, m. \end{aligned}$$

Often useful trick: **variable splitting**

$$\begin{aligned} \min_{x_1,\ldots,x_T,z} \quad & \sum_{i=1}^{T} f_{0,i}(x_i) \\ \text{s.t.} \quad & f_i(z) \le 0, \quad i = 1, \ldots, m \\ & x_i = z, \quad i = 1, \ldots, T. \end{aligned}$$

Introducing equality constraints

Separable function

$$\begin{aligned} \min \quad & \sum_{i=1}^{T} f_{0,i}(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Often useful trick: **variable splitting**

$$\begin{aligned} \min_{x_1,\ldots,x_T,z} \quad & \sum_{i=1}^{T} f_{0,i}(x_i) \\ \text{s.t.} \quad & f_i(z) \leq 0, \quad i = 1, \ldots, m \\ & x_i = z, \quad i = 1, \ldots, T. \end{aligned}$$

Almost separate problems! Useful for distributed computing.

Constraint removal

Constrained problem

$$\min \quad f_0(x) \quad \text{s.t.} \quad x \in \mathcal{X}.$$

Constraint removal

Constrained problem

$$\min \quad f_0(x) \quad \text{s.t.} \quad x \in \mathcal{X}.$$

Unconstrained problem

$$\min \quad f_0(x) + \mathbb{I}_X(x)$$

Constraint removal

Constrained problem

$$\min \quad f_0(x) \quad \text{s.t.} \quad x \in \mathcal{X}.$$

Unconstrained problem

$$\min \quad f_0(x) + \mathbb{I}_X(x)$$

Penalized form (approximate)

$$\min \quad f_0(x) + \rho \|\max\{0, f(x)\}\|_2^2,$$

where $f(x) = [f_1(x), \ldots, f_m(x)]^T$; $\rho \gg 0$.

Constraint removal

Constrained problem

$$\min \quad f_0(x) \quad \text{s.t.} \quad x \in \mathcal{X}.$$

Unconstrained problem

$$\min \quad f_0(x) + \mathbb{I}_X(x)$$

Penalized form (approximate)

$$\min \quad f_0(x) + \rho \|\max\{0, f(x)\}\|_2^2,$$

where $f(x) = [f_1(x), \ldots, f_m(x)]^T$; $\rho \gg 0$.

Reducing number of constraints

$$\begin{aligned} &\min f_0(x) \quad \text{s.t.} \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\ \Longrightarrow &\min f_0(x) \quad \text{s.t.} \quad [g(x) := \max_{1 \leq i \leq m} f_i(x)] \leq 0. \end{aligned}$$

Implicit constraints

$$\min \quad c^T x - \sum_{i=1}^{m} \log(b_i - a_i^T x),$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and a_i^T are rows of $A \in \mathbb{R}^{m \times n}$.

Implicit constraints

$$\min \quad c^T x - \sum_{i=1}^{m} \log(b_i - a_i^T x),$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and a_i^T are rows of $A \in \mathbb{R}^{m \times n}$.

- Implicit constraint: $b_i - a_i^T x \in \mathrm{dom}(-\log z)$ (else $+\infty$)
- Thus, x must be in strict interior of $\mathcal{P} = \{x \mid Ax \leq b\}$.

Implicit constraints

$$\min \quad c^T x - \sum_{i=1}^m \log(b_i - a_i^T x),$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and a_i^T are rows of $A \in \mathbb{R}^{m \times n}$.

- Implicit constraint: $b_i - a_i^T x \in \operatorname{dom}(-\log z)$ (else $+\infty$)
- Thus, x must be in strict interior of $\mathcal{P} = \{x \mid Ax \le b\}$.

$\boxed{\text{Idea comes up again in interior point methods}}$

Problem classes

Linear Programming

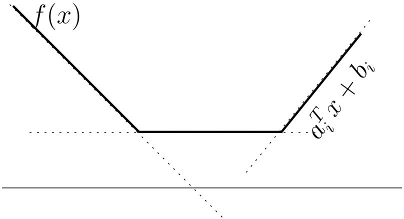
$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \le b, \quad Cx = d. \end{aligned}$$

Linear Programming

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \le b, \quad Cx = d. \end{aligned}$$

Piecewise linear minimization

$$\min \quad f(x) = \max_{1 \le i \le m} (a_i^T x + b_i)$$

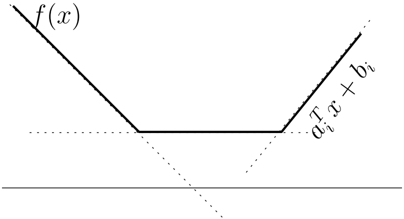


Linear Programming

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \le b, \quad Cx = d. \end{aligned}$$

Piecewise linear minimization

$$\min \quad f(x) = \max_{1 \le i \le m} (a_i^T x + b_i)$$



$$\min_{x,t} \quad t \qquad \text{s.t.} \quad a_i^T x + b_i \le t, \quad i = 1, \ldots, m.$$

- Linear program with variables $x, t \in \mathbb{R}$.

LP Exercises

- Formulate $\min \|Ax - b\|_1$ as an LP ($\|x\|_1 = \sum_i |x_|$)
- Formulate $\min \|Ax - b\|_\infty$ as an LP ($\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$)

Quadratic Programming

$$\min \quad \tfrac{1}{2}x^T Ax + b^T x + c \qquad \text{s.t.} \;\; Gx \leq h.$$

We assume $A \succeq 0$ (semidefinite).

Quadratic Programming

$$\min \quad \tfrac{1}{2}x^T Ax + b^T x + c \qquad \text{s.t.} \quad Gx \leq h.$$

We assume $A \succeq 0$ (semidefinite).

Exercise: Say no constraints; does this QP always have a solution?

Quadratic Programming

$$\min \quad \tfrac{1}{2}x^T A x + b^T x + c \qquad \text{s.t. } \; Gx \leq h.$$

We assume $A \succeq 0$ (semidefinite).
Exercise: Say no constraints; does this QP always have a solution?

Nonnegative least squares (NNLS)

$$\min \quad \tfrac{1}{2}\|Ax - b\|^2 \quad \text{s.t. } x \geq 0.$$

Exercise: Prove that NNLS always has a solution.

Regularized least-squares

Lasso

$$\min \quad \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_1.$$

Exercise: How large must $\lambda > 0$ so that $x = 0$ is the optimum?

Regularized least-squares

Lasso

$$\min \quad \tfrac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_1.$$

Exercise: How large must $\lambda > 0$ so that $x = 0$ is the optimum?

Total-variation denoising

$$\min \quad \tfrac{1}{2}\|Ax - b\|_2^2 + \lambda \sum_{i=1}^{n-1} |x_{i+1} - x_i|.$$

Exercise: Is the total-variation term a norm? Prove or disprove.

Regularized least-squares

Lasso

$$\min \quad \tfrac{1}{2}\|Ax-b\|_2^2 + \lambda\|x\|_1.$$

Exercise: How large must $\lambda > 0$ so that $x = 0$ is the optimum?

Total-variation denoising

$$\min \quad \tfrac{1}{2}\|Ax-b\|_2^2 + \lambda \sum_{i=1}^{n-1} |x_{i+1} - x_i|.$$

Exercise: Is the total-variation term a norm? Prove or disprove.

Group Lasso

$$\min_{x_1,\ldots,x_T} \quad \frac{1}{2}\left\| b - \sum_{j=1}^{T} A_j x_j \right\|_2^2 + \lambda \sum_{j=1}^{T} \|x_j\|_2.$$

$\boxed{\text{Notice non-differentiable regularizers}}$

Second order cone program (SOCP)

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

Second order cone program (SOCP)

$$\begin{aligned} \min \quad & f^T x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m. \end{aligned}$$

- Linear objective
- Nonlinear, nondifferentiable constraints
- Generalization of LP, QP: allows cone constraints
- Recall $\mathcal{Q}^n := \left\{(x, t) \in \mathbb{R}^{n+1} \mid \|x\|_2 \leq t\right\}$ is a convex cone

Example – robust LP

$$\min \quad c^T x, \qquad \text{s.t.} \;\; a_i^T x \leq b_i \; \forall a_i \in \mathcal{E}_i$$
$$\mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}$$

The constraints are uncertain but with bounded uncertainty.

Example – robust LP

$$\min \quad c^T x, \qquad \text{s.t. } \; a_i^T x \le b_i \;\forall a_i \in \mathcal{E}_i$$
$$\mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}$$

The constraints are uncertain but with bounded uncertainty.

$$\min_x \sup_{\|u\|_2 \le 1} \left\{ c^T x \mid a_i^T x \le b_i, \quad a_i \in \mathcal{E}_i \right\}$$

Example – robust LP

$$\min \quad c^T x, \qquad \text{s.t.} \;\; a_i^T x \le b_i \;\forall a_i \in \mathcal{E}_i$$
$$\mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}$$

The constraints are uncertain but with bounded uncertainty.

$$\min_x \sup_{\|u\|_2 \le 1} \left\{ c^T x \mid a_i^T x \le b_i, \quad a_i \in \mathcal{E}_i \right\}$$

SOCP formulation

$$\min \quad c^T x, \qquad \text{s.t.} \;\; \|P_i^T x\|_2 \le -\bar{a}_i^T x + b_i, i = 1, \dots, m.$$

Example – robust LP

$$\min \quad c^T x, \qquad \text{s.t. } a_i^T x \le b_i \ \forall a_i \in \mathcal{E}_i$$
$$\mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}$$

The constraints are uncertain but with bounded uncertainty.

$$\min_x \sup_{\|u\|_2 \le 1} \left\{ c^T x \mid a_i^T x \le b_i, \quad a_i \in \mathcal{E}_i \right\}$$

SOCP formulation

$$\min \quad c^T x, \qquad \text{s.t. } \|P_i^T x\|_2 \le -\bar{a}_i^T x + b_i, i = 1, \ldots, m.$$

Semidefinite Program (SDP)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & A(x) := A_0 + x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq 0. \end{aligned}$$

Semidefinite Program (SDP)

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad A(x) := A_0 + x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq 0.$$

- $A_0, \ldots, A_n$ are real, symmetric matrices
- Inequality $A \preceq B$ means $B - A$ is *semidefinite*
- Also a cone program (conic optimization problem)

Semidefinite Program (SDP)

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad A(x) := A_0 + x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq 0.$$

- $A_0, \ldots, A_n$ are real, symmetric matrices
- Inequality $A \preceq B$ means $B - A$ is *semidefinite*
- Also a cone program (conic optimization problem)
- SDP $\supset$ SOCP $\supset$ QP $\supset$ LP
- **Exercise:** Write LPs, QPs, and SOCPs as SDPs

Semidefinite Program (SDP)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & A(x) := A_0 + x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq 0. \end{aligned}$$

- $A_0, \ldots, A_n$ are real, symmetric matrices
- Inequality $A \preceq B$ means $B - A$ is *semidefinite*
- Also a cone program (conic optimization problem)
- SDP $\supset$ SOCP $\supset$ QP $\supset$ LP
- **Exercise:** Write LPs, QPs, and SOCPs as SDPs
- Feasible set of SDP is $\{$semidefinite cone $\bigcap$ hyperplanes$\}$

Semidefinite Program (SDP)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & A(x) := A_0 + x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq 0. \end{aligned}$$

- $A_0, \ldots, A_n$ are real, symmetric matrices
- Inequality $A \preceq B$ means $B - A$ is *semidefinite*
- Also a cone program (conic optimization problem)
- SDP $\supset$ SOCP $\supset$ QP $\supset$ LP
- **Exercise:** Write LPs, QPs, and SOCPs as SDPs
- Feasible set of SDP is $\{$semidefinite cone $\bigcap$ hyperplanes$\}$
- When is a convex problem **representable** as an SDP?

Examples

♠ **Eigenvalue optimization:** $\min \lambda_{\max}(A(x))$

$$\min \quad t \quad \text{s.t.} \quad A(x) \preceq tI.$$

Examples

♠ **Eigenvalue optimization:** $\min \lambda_{\max}(A(x))$

$$\min \quad t \quad \text{s.t.} \quad A(x) \preceq tI.$$

♠ **Norm minimization:** $\min \|A(x)\|$

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

Examples

♠ **Eigenvalue optimization:** $\min \lambda_{\max}(A(x))$

$$\min \quad t \quad \text{s.t.} \quad A(x) \preceq tI.$$

♠ **Norm minimization:** $\min \|A(x)\|$

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

♠ Many more examples! See CVX documentation also.

♠ SDP relaxations of nonconvex problems — powerful, important

♠ More on this next lecture

References

1. L. Vandenberghe. MLSS 2012 Lecture slides.