EECS 127/227AT Optimization Models in Engineering
Spring 2020 Discussion 11

1. Dual norms and SOCP

Consider the problem

$$p^* = \min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{y}\|_1 + \mu\|\vec{x}\|_2,$$

$$\min \vec{c}^\top \vec{x}$$

$$\|A\vec{x} + \vec{b}\|_2 \leq \vec{d}^\top \vec{x} + e$$

$$\dots$$

where $A \in \mathbb{R}^{m\times n}$, $\vec{y} \in \mathbb{R}^m$, and $\mu > 0$.

$$\sum_i |(A\vec{x} - \vec{y})_i| = \sum_i |\vec{a}_i^\top \vec{x} - y_i|$$

One row of A $\quad A = \begin{pmatrix} \vec{a}_1^\top \\ \vdots \\ \vec{a}_m^\top \end{pmatrix}$

Should be able to do

(a) Express this (primal) problem in standard SOCP form.

$$p^* = \min_{\vec{x}, \vec{t}} \underbrace{\sum_i t_i}_{= \vec{1}^\top \vec{t}} + \mu \|\vec{x}\|_2$$

$$t_i \geq |\vec{a}_i^\top \vec{x} - y_i| \quad \forall i$$

SOCP

$$p^* = \min_{\vec{x}, \vec{t}, z} \vec{1}^\top \vec{t} + \mu z \quad \leftarrow \text{linear}$$

$$-\vec{a}_i^\top \vec{x} + y_i \leq t_i \quad \forall i \in [\![1, m]\!] \quad \leftarrow \text{linear}$$

$$\vec{a}_i^\top \vec{x} - y_i \leq t_i \quad \forall i \in [\![1, m]\!] \quad \leftarrow \text{linear}$$

$$\|\vec{x}\|_2 \leq z \quad \leftarrow \text{conic}$$

Might need def of dual prob here

(b) Find a dual to the problem and express it in standard SOCP form.
Hint: Recall that for every vector $\vec{z}$, the following dual norm equalities hold:

$$\|\vec{z}\|_2 = \max_{\vec{u}\,:\,\|\vec{u}\|_2 \leq 1} \vec{u}^\top \vec{z}, \quad \|\vec{z}\|_1 = \max_{\vec{u}\,:\,\|\vec{u}\|_\infty \leq 1} \vec{u}^\top \vec{z}.$$

How to define dual when there is no constraints?

$$\|A\vec{x} - \vec{y}\|_1 = \max_{\vec{u} : \|\vec{u}\|_\infty \leq 1} \vec{u}^\top (A\vec{x} - \vec{y})$$

Should be given to you

$$\|\vec{x}\|_2 = \max_{\vec{v} : \|\vec{v}\|_2 \leq 1} \vec{v}^\top \vec{x}$$

$$p^* = \min_{\vec{x}} \max_{\substack{\vec{u}, \vec{v} \\ \|\vec{u}\|_\infty \leq 1 \\ \|\vec{v}\|_2 \leq 1}} \underbrace{\vec{u}^\top (A\vec{x} - \vec{y}) + \vec{v}^\top \mu \vec{x}}_{\mathcal{L}(\vec{x}, \vec{u}, \vec{v})}$$

continuous and diff w.r.t. $\vec{x}$

linear w.r.t. $\vec{x}$

min $\vec{c}^\top \vec{x}$ $= -\infty$ if $\vec{c} \neq 0$
$=$ cste if $\vec{c} = \vec{0}$

$$d^* = \max_{\substack{\vec{u}, \vec{v} \\ \|\vec{u}\|_\infty \leq 1 \\ \|\vec{v}\|_2 \leq 1}} \boxed{\min_{\vec{x}} \vec{u}^\top (A\vec{x} - \vec{y}) + \mu \vec{v}^\top \vec{x}}$$

$$\nabla_{\vec{x}} \mathcal{L}(\vec{x}, \vec{u}, \vec{v}) = A^\top \vec{u} + \mu \vec{v} = \vec{0} \text{ if } x^* \min$$

$$d^* = \max_{\substack{\vec{u}, \vec{v} \\ \|\vec{u}\|_\infty \le 1,\ \|\vec{v}\|_2 \le 1}} \begin{cases} -\vec{u}^T\vec{y} & \text{if } \underbrace{A^T\vec{u} + \mu\vec{v} = 0} \\ -\infty & \text{otherwise} \end{cases}$$

Should have some intuition about that

(c) Assume strong duality holds[1] and that $m = 100$ and $n = 10^6$, i.e., A is 100×10^6. Which problem would you choose to solve using a numerical solver: the primal or the dual? Justify your answer.

Primal variable is in $\mathbb{R}^n$

Dual variable is in $\mathbb{R}^m$ SOCP

$$d^* = \max_{\vec{u}, \vec{v}} \ -\vec{u}^T\vec{y}$$
$$\|\vec{u}\|_\infty \le 1$$
$$\|\vec{v}\|_2 \le 1$$
$$A^T\vec{u} + \mu\vec{v} = 0 \ \Big\} \to \|A^T\vec{u}\|_2 \le \mu$$

$$d^* = \max_{\vec{u}} \ -\vec{u}^T\vec{y}$$
$$\|\vec{u}\|_\infty \le 1$$
$$\|A^T\vec{u}\|_2 \le \mu$$

Sometime duality will transpose your data matrix and might reduce the dimensionality of your problem

2. Squaring SOCP constraints — Prevent you to do this error and understand why it is an error

When considering a second-order cone (SOC) constraint, you might be tempted to square it to obtain a classical convex quadratic constraint. This problem explores why that might not always work, and how to introduce additional constraints to maintain equivalence and convexity.

(a) For $\vec{x} \in \mathbb{R}^2$, consider the constraint

$x_1 - 2x_2 \ge 0$ Why squaring? SOCP

Convex because SOC $\longrightarrow$
$$x_1 - 2x_2 \ge \|\vec{x}\|_2,$$

$$\|A\vec{x} + \vec{b}\|_2 \le \underbrace{\vec{c}^T\vec{x} + d}$$

Cannot square

and its squared counterpart

Not convex $\longrightarrow$
$$(x_1 - 2x_2)^2 \ge \|\vec{x}\|_2^2.$$

$$\vec{x}^TA^TA\vec{x} + 2\vec{b}^TA\vec{x} + \vec{b}^T\vec{b}$$

QP? $\to$ Some time not a QP

$$\le (\vec{c}^T\vec{x} + d)(\vec{c}^T\vec{x} + d)$$
$$\vec{x}^T\vec{c}\vec{c}^T\vec{x}$$
$$\vec{x}^T\underbrace{(A^TA - \vec{c}\vec{c}^T)}\vec{x} + \dots \le 0$$

Might not be PSD

Are the two sets equivalent? Are they both convex?

No

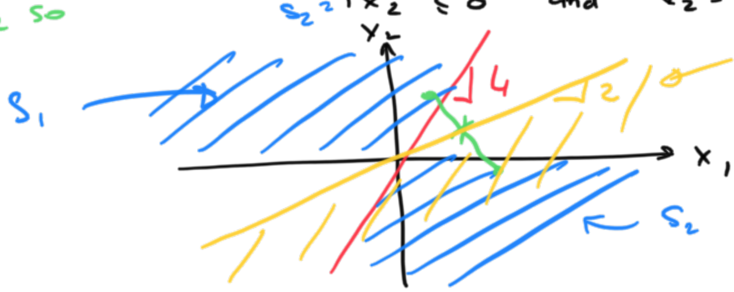
$$\cancel{x_1^2} + \cancel{2}x_2^2 - 4x_1x_2 \ge \cancel{x_1^2} + \cancel{x_2^2}$$

Not convex constraint $\longrightarrow$ $x_2(x_2 - 4x_1) \ge 0$

It is quadratic but not positive so

$S_1 = \{x_2 \ge 0$ and $x_2 - 4x_1 \ge 0\}$

or $S_2 = \{x_2 \le 0$ and $x_2 - 4x_1 < 0\}$



$S_1 \cup S_2 = \{\vec{x},\ (x_1 - 2x_2)^2 \ge \|\vec{x}\|_2^2\}$

$S_2 = \{\vec{x},\ x_1 - 2x_2 \ge \|\vec{x}\|_2\}$

[1] In fact, you can show that strong duality holds using Sion's theorem, a generalization of the minimax theorem that is beyond the scope of this class.

(b) What additional constraint must be imposed alongside the squared constraint to enforce the same feasible set as the unsquared SOC constraint?

↳ In $x_1 - 2x_2 \geq \|\vec{x}\|_2$

If you want to square the inequality you should add the constraint $x_1 - 2x_2 \geq 0$

$$\|A\vec{x} + \vec{b}\|_2 \leq \vec{c}^\top \vec{x} + d \iff \|A\vec{x} + \vec{b}\|_2^2 \leq \vec{c}^\top \vec{x} + d \text{ and } 0 \leq \vec{c}^\top \vec{x} + d$$

SOC

$$\vec{x}^\top A^\top A \vec{x} + 2\vec{b}^\top A \vec{x} + \vec{b}^\top \vec{b} \leq \vec{x}^\top \vec{c}\vec{c}^\top \vec{x} + 2d\vec{c}^\top \vec{x} + d^2$$

$$\iff \vec{x}^\top \underbrace{(A^\top A - c^\top c)}_{\text{might not be PSD}} \vec{x} + 2(A^\top \vec{b} - d\vec{c})^\top \vec{x} + \vec{b}^\top \vec{b} - d^2 \leq 0 \text{ and } 0 \leq \vec{c}^\top \vec{x} + d$$

~~CP~~

LP

3. Casting optimization problems as SOCPs

Cast the following problem as an SOCP in its standard form:

$$\begin{cases} \min\limits_{\vec{x} \in \mathbb{R}^n} & \sum\limits_{i=1}^{p} \dfrac{\|F_i \vec{x} + \vec{g}_i\|_2^2}{\vec{a}_i^\top \vec{x} + b_i} \\ \text{s.t.} & \vec{a}_i^\top \vec{x} + b_i > 0, \quad i = 1, \ldots, p, \end{cases}$$

where $F_i \in \mathbb{R}^{m \times n}$, $\vec{g}_i \in \mathbb{R}^m$, $\vec{a}_i \in \mathbb{R}^n$, and $b_i \in \mathbb{R}$, for $i = 1, \ldots, p$.

Intuition, practice, linear algebra, duality

$$\min_{\vec{x}, \vec{t}} \quad \vec{1}^\top \vec{t} \quad \leftarrow \text{Linear}$$

$$\frac{\|F_i \vec{x} + \vec{g}_i\|_2^2}{\vec{a}_i^\top \vec{x} + b_i} \leq t_i \quad \forall i$$

$$\vec{a}_i^\top \vec{x} + b_i > 0 \quad \forall i \leftarrow \text{Linear}$$

$$\|F_i \vec{x} + \vec{g}_i\|_2^2 \leq t_i (\vec{a}_i^\top \vec{x} + b_i) \quad \forall i$$

$$\left\| \begin{matrix} 2(F_i \vec{x} + \vec{g}_i) \\ t_i - \vec{a}_i^\top \vec{x} - b_i \end{matrix} \right\|_2 \leq t_i + \vec{a}_i^\top \vec{x} + b_i \leftarrow \text{SOC}$$

SOCP

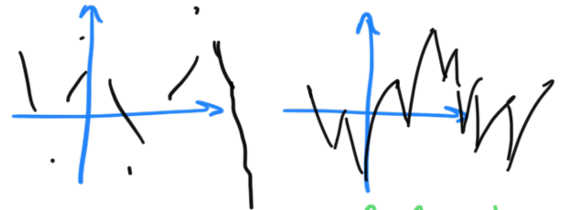
4. A review of standard problem formulations

In this question, we review conceptually the standard forms of various problems and the assertions we can (and cannot!) make about each.

(a) ***Linear programming (LP).***

i. Write the most general form of a linear program (LP) and list its defining attributes.

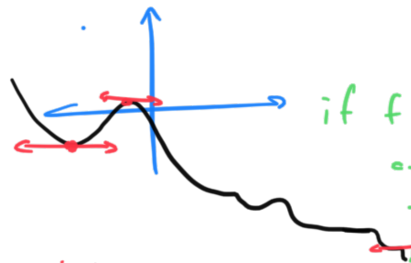
Convex problem : $\min_{\vec{x} \in \mathbb{R}} f(\vec{x})$



if f not continuous almost impossible to minimize it

if f continuous but not diff → you might find local minimum hard

Small analysis shows that we like problem without constraints, convex, continuous and differentiable



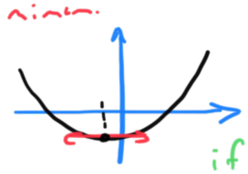
if f continuous and diff → set $\nabla f = 0$ to find local extreme or inflexion point

ii. Under what conditions is an LP convex?

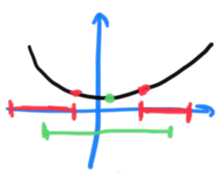
If we add constraints :

$\min_{\vec{x}} f(\vec{x})$ ← continuous, convex, has a minimum, differentiable

$\vec{x} \in X$ ← feasible set

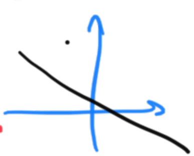


if f continuous, diff and convex → set $\nabla f = 0$ to get global minimum if it exist



if X is not convex the problem is hard to solve

We like x to be convex to



$\min_{\vec{x}} f(\vec{x})$ ← continuous, diff, convex, has min

$\vec{x} \in X$ ← set is convex

$y = -x$ is convex and does not have global minimum

(b) ***Quadratic programming (QP).***

i. Write the most general form of a quadratic program (QP) and list its defining attributes.

$\min_{\vec{x}} f(\vec{x})$ ← continuous, diff, has a min, convex

$f_i(\vec{x}) \leq 0 \quad \forall i = 1, \dots, n$ ← convex ← sublevel set of convex function is convex

$h_j(\vec{x}) = 0 \quad \forall j = 1, \dots, m$ ← affine ← level set of affine function is convex (because affine)

Define a convex problem

Define convex set of feasible point

we like that : easy (easier) to solve

With constraints : you cannot find min with $\nabla f = 0$

∘ You can still do some kind of algorithms

$$\min_{\bar{x}} f(\bar{x}) \quad \bar{x} \in X \xrightarrow{\text{duality}} = \min_{\bar{x}} \tilde{f}(\bar{x}) \qquad \tilde{f}(\bar{x}) = \begin{cases} f(x) & \text{if } \bar{x} \in X \\ 0 & \text{otherwise} \end{cases}$$

$$= \max_{\bar{\lambda} \geq 0} \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu})$$

~~ii. Under what conditions is a QP convex?~~

If $\tilde{f}$ is convex, continuous, diff and min exist, can be solved by $\nabla f = 0$

Stationary condition of KKT cond.

$$\nabla_{\bar{x}} \mathcal{L}(\bar{x}^*, \bar{\lambda}^*, \bar{\mu}^*) = 0$$

All theory you should know.

Applications on how to solve pbs efficiently → Algorithms GD, type of problems LP, QP, QCQP, SOCP, SDP interior point methods

how to use it in engineering

→ Regression, LASSO, SVM, Machine learning,

(c) ***Quadratically-constrained quadratic programming (QCQP).***

i. Write the most general form of a quadratically-constrained quadratic program (QCQP) and list its defining attributes.

pb 1 dis ← complexity reduction,

pb 1 dis 10 ← pb equivalence

ii. Under what conditions is a QCQP convex?

(d) ***Second-order cone programming (SOCP).***

i. Write the most general form of a second-order cone program (SOCP) and list its defining attributes.

ii. Under what conditions is an SOCP convex?

(e) ***Relationships.*** Recall that

$$LP \subset QP_{\text{convex}} \subset QCQP_{\text{convex}} \subset SOCP \subset \{\text{all convex programs}\},$$

where LP denotes the set of all linear programs, QP_{convex} denotes the set of all convex quadratic programs, etc. Which of these problems can be solved most efficiently? Why are these categorizations useful?