Convex and Nonsmooth Optimization Homework 2

Xiaoou (Charlotte) Cheng

February 8, 2022

1. Suppose a function f is convex. For two points $(x_1, t_1), (x_2, t_2)$ in its epigraph, we have, for any $\alpha \in [0, 1]$,

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) \le \alpha t_1 + (1-\alpha)t_2.$$

Thus, $(\alpha x_1 + (1-\alpha)x_2, (\alpha t_1 + (1-\alpha)t_2)$ is in the epigraph of f, which shows that the epigraph is a convex set.

Suppose the epigraph of a function f is a convex set. We take the conventional definition of a convex function here. Then, for any $x_1, x_2 \in \mathbf{dom} f$, we know that $(x_1, f(x_1)), (x_2, f(x_2)) \in \mathbf{epi} f$, and by convexity of the epigraph, for any $\alpha \in [0, 1]$

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2).$$

We also need to show that $\mathbf{dom} f$ is a convex set, which is true because $\mathbf{epi} f = \{(x, t) | x \in \mathbf{dom} f, f(x) \le t\}$ is a convex set, and $\mathbf{dom} f$ is $\mathbf{epi} f$ projected to the first coordinate. The projection is an affine transformation (and it is actually linear), and the image of a convex set under an affine function is convex, so $\mathbf{dom} f$ is a convex set.

2. (a) Suppose $y_1, y_2 \in K^*$. For any $t_1, t_2 \ge 0$, we have, for any $x \in K$,

$$x^T(t_1 y_1 + t_2 y_2) = t_1 x^T y_1 + t_2 x^T y_2 \ge 0,$$

so $t_1 y_1 + t_2 y_2 \in K^*$ and K^* is a convex cone.

(b) For any $y \in K_2^*$, we have

$$x^T y \ge 0, \forall x \in K_2.$$

Since $K_1 \subseteq K_2$, $x^T y \ge 0, \forall x \in K_1$, which shows that $y \in K_1^*$. Since y is arbitrary in K_2^*, we have shown that $K_2^* \subseteq K_1^*$.

3. By the definition of the dual cone, $y \in \mathbb{R}^m$ is in the dual cone if and only if

$$x^T A^T y \ge 0, \forall x \succeq 0. \tag{1}$$

We know that the cone $\mathbb{R}^n_+$ is self-dual. Thus, (1) is satisfied if and only if $A^T y \ge 0$. Therefore, the dual cone is

$$\left\{y \in \mathbb{R}^m | A^T y \succeq 0\right\}.$$

4. For any $(x, t), (y, s) \in K$, by Cauchy-Schwarz inequality,

$$|x^T y| \le \|x\|_2 \|y\|_2 \le ts \quad \Longrightarrow x^T y \ge -ts.$$

Thus, $x^T y + ts \ge 0$, and since (y, s) is an arbitrary element in K, we have $(x, t) \in K^*$, which implies that $K \subseteq K^*$.

On the other hand, suppose $(x, t) \in K^*$ but $(x, t) \notin K$, then $\|x\|_2 > t$. Since (by Cauchy-Schwarz inequality)

$$\|x\|_2 = \sup_{z: \|z\|_2 \le 1} x^T z,$$

there exists z such that

$$\|z\|_2 \le 1, \quad x^T z > t.$$

Let $y = -sz$ for some $s > 0$. $\|y\|_2 \le s$, so $(y, s) \in K$, but

$$x^T y + ts = -sx^T z + ts < 0,$$

which contradicts with the assumption that $(x, t) \in K^*$. This shows that $K^* \subseteq K$, and thus, $K^* = K$.

5. (a) The conjugate function is

$$f^*(y) = \begin{cases} 0, & \text{if } y \ge 0, \mathbb{1}^T y = 1, \\ \infty, & \text{otherwise.} \end{cases}$$

Here is the analysis. If there exists a $y_i < 0$, then we take x such that $x_j = 0$ for all $j \ne i$, and let $x_i \to -\infty$. Then,

$$y^T x - f(x) = y_i x_i \to +\infty \Longrightarrow f^*(y) = +\infty.$$

If there exists a $y_i > 1$, then we take x such that $x_j = 0$ for all $j \ne i$, and let $x_i \to +\infty$. Then,

$$y^T x - f(x) = y_i x_i - x_i = (y_i - 1) x_i \to +\infty \Longrightarrow f^*(y) = +\infty.$$

Now we consider y such that $0 \le y_i \le 1$ for all i. Given any x, suppose $x_i = \max_k x_k$, then

$$y^T x - f(x) = y^T x - x_i \le y^T(\mathbb{1} x_i) - x_i = x_i(y^T \mathbb{1} - 1),$$

since $y_i \ge 0$ and the equality is attained if $x_j = x_i$ for all $1 \le j \le n$.
If $y^T \mathbb{1} \ne 1$, then we can take $x_i \to +\infty$ or $-\infty$ so $y^T x - f(x) \to +\infty$. In other words,

$$f^*(y) = +\infty \quad \text{if } 0 \le y_i \le 1, y^T \mathbb{1} \ne 1.$$

If $y^T \mathbb{1} = 1$, then following the preceding analysis, we have

$$y^T x - f(x) \le y^T(\mathbb{1} x_i) - x_i = 0,$$

which is attained if all x_i's are equal.
Thus,

$$f^*(y) = 0 \quad \text{if } 0 \le y_i \le 1, y^T \mathbb{1} = 1.$$

(b) The conjugate function is

$$f^*(y) = \begin{cases} 0, & \text{if } y \in C, \\ +\infty, & \text{otherwise.} \end{cases}$$

with

$$C = \left\{ y \in \mathbb{R}^n | 0 \le y \le 1, y_{[j]} + \dots y_{[n]} \ge r + 1 - j, \forall 2 \le j \le r, \sum_{k=1}^n y_k = r \right\}.$$

In fact[1], since $0 \le y_i \le 1$, C can be simplified to

$$C = \left\{ y \in \mathbb{R}^n | 0 \le y \le 1, \sum_{k=1}^n y_k = r \right\}.$$

Here is the analysis.
If there is a $y_i > 1$, then let x such that $x_j = 0$ for $j \ne i$ and $x_i \to +\infty$, and we know that $f^*(y) = +\infty$.
Similarly, if there is a $y_i < 0$, then let x such that $x_j = 0$ for $j \ne i$ and $x_i \to -\infty$, and we know that $f^*(y) = +\infty$.

[1]Upon looking at this SE post afterwards, I realized that my expression can be simplified

Now suppose $0 \le y \le 1$. If $y \notin C$, then there exists some $y_{i_1} + \cdots + y_{i_{n-j+1}} < r + 1 - j$ for some $2 \le j \le r$, or it might be that $\sum_{k=1}^n y_k \ne r$.

Consider the first case, where we have $y_{i_1} + \cdots + y_{i_{n-j+1}} < r + 1 - j$ for some $2 \le j \le r$. Then we can take $x_{i_1} = x_{i_2} = \cdots = x_{i_{n-j+1}} \to -\infty$ and let all other $x_k = 0$. Thus,

$$y^T x - f(x) = (y_{i_1} + \dots y_{i_{n-j+1}})x_{i_1} - (r - j + 1)x_{i_1} = (y_{i_1} + \dots y_{i_{n-j+1}} - (r - j + 1))x_{i_1} \to +\infty.$$

Therefore,

$$f^*(y) = +\infty.$$

Now suppose $0 \le y \le 1$, $y_{[j]} + \dots y_{[n]} \ge r + 1 - j, \forall 2 \le j \le r$, but $\sum_{k=1}^n y_k \ne r$. If $y^T \mathbb{1} > r$, then take $x_j = x_1$ for all j, we have

$$y^T x - f(x) = y^T \mathbb{1} x_1 - r x_1 \to +\infty$$

as $x_1 \to +\infty$, so

$$f^*(y) = +\infty.$$

If $y^T \mathbb{1} < r$, then let $x_j = x_1$ for all j, we have, as $x_1 \to -\infty$,

$$y^T x - f(x) = y^T \mathbb{1} x_1 - r x_1 \to +\infty.$$

So

$$f^*(y) = +\infty.$$

Therefore, if $y \notin C$, $f^*(y) = +\infty$.

Now we consider the case where $y \in C$. Given any $x \in \mathbb{R}^n$, we denote $x_{i_j} = x_{[j]}$. Since $y \in C$, we have

$$\begin{aligned}
y^T x - f(x) &\le y_{i_1} x_{i_1} + \cdots + y_{i_{r-1}} x_{i_{r-1}} + (y_{i_r} + \cdots + y_{i_n} - 1)x_{i_r} - (x_{i_1} + \cdots + x_{i_{r-1}}) \\
&\le y_{i_1} x_{i_1} + \cdots + y_{i_{r-2}} x_{i_{r-2}} + (y_{i_{r-1}} + y_{i_r} + \cdots + y_{i_n} - 2)x_{i_{r-1}} - (x_{i_1} + \cdots + x_{i_{r-2}}) \\
&\le \dots \\
&\le (y_{i_1} + \cdots + y_{i_n} - r)x_{i_1} \\
&= (y^T \mathbb{1})x_{i_1} \\
&= 0.
\end{aligned}$$

Thus, $f^*(y) \le 0$. On the other hand, we can take x such that $x_i = 0$ for all i to get $y^T x - f(x) = 0$. So

$$f^*(y) = 0, \quad \forall y \in C.$$

6. The code is listed below. Here in Fig 1, we rerun the example on the webpage with our function with `A` and `b` and the result is on the left, which agrees with the true result and shows the reliability of our function. On the right, we add two other constraints, which are

$$-2x - y \le 1,$$
$$4x - y \le 1.$$

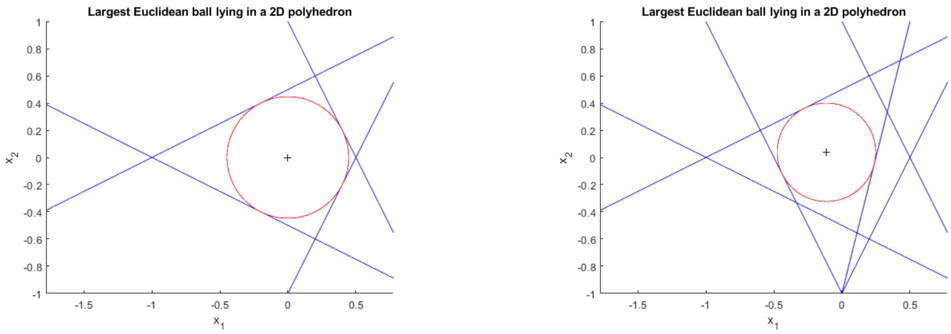

Figure 1: Left: reproduce the example on the webpage; Right: add two additional constraints

```
function [x_c, r] = ChebyCenter(A,b)
% each column of A corresponds to a line
% --- input ---
% A: a matrix; b: a vector -- the RHS
% A(:,k) ... b(k)
% --- output ---
% x_c: the location of the center
% r: the radius

n = size(A,2);
B = A';
len = sqrt(sum(B.^2 ,2));


% Create and solve the model
cvx_begin
    variable r(1)
    variable x_c(2)
    maximize ( r )
    B*x_c + r* len ≤ b;
cvx_end

% Generate the figure
x = linspace(-2,2);
theta = 0:pi/100:2*pi;

hold on

for i = 1:n
    plot( x, -x*A(1,i)./A(2,i) + b(i)./A(2,i),'b-');
end
plot( x_c(1) + r*cos(theta), x_c(2) + r*sin(theta), 'r');
plot(x_c(1),x_c(2),'k+')
xlabel('x_1')
ylabel('x_2')
title('Largest Euclidean ball lying in a 2D polyhedron');
axis([-1 1 -1 1])
axis equal
```

The main function for this question:

```
close all;
clear variables;

%A = [2,2,-1,-1;1,-1,2,-2]; % original
A = [2,2,-1,-1,-2, 4; 1,-1,2,-2,-1,-1];
 %b = ones(4,1); % original
b = ones(6,1);
[x_c, r] = ChebyCenter(A,b);
```

7. If there is no point inside the polyhedron, the output `r` is nonpositive. In the plot, we can get a circle, though. This is because the way we plot the result includes `r * cos(theta)` and `r * sin(theta)` which are well-defined with any r. But since the polehedron is empty, we shouldn't expect to get any meaningful result. The function doesn't pop out an error because we don't require $r \geq 0$ in our LP.

   Here is an example where we add the following additional constraint to the original problem

$$2x + y \leq -3.$$

   The result is as follows, which gives `r=-0.149071`.

```
close all;
clear variables;
```

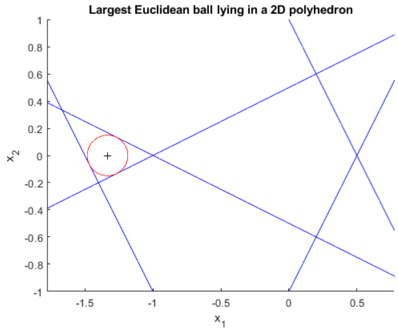

Figure 2: Polyhedron doesn't contain any point

```

A = [2,2,-1,-1,2; 1,-1,2,-2,1];
b = [1;1;1;1;-3];

[x_c, r] = ChebyCenter(A,b);
```

8. Following the book, now we require that

$$\|u\|_p \leq r \Longrightarrow a_i^T(x_c + u) \leq b_i.$$

Since

$$\sup\left\{a_i^T u \middle| \|u\|_p \leq r\right\} = r\|a_i\|_q,$$

where q is conjugate to p in the sense that $1/p + 1/q = 1$ (p or q can be ∞), we can write the condition as

$$a^T x_c + r\|a_i\|_q \leq b_i,$$

which is a linear inequality in x_c and r.

The optimization problem is an LP as

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

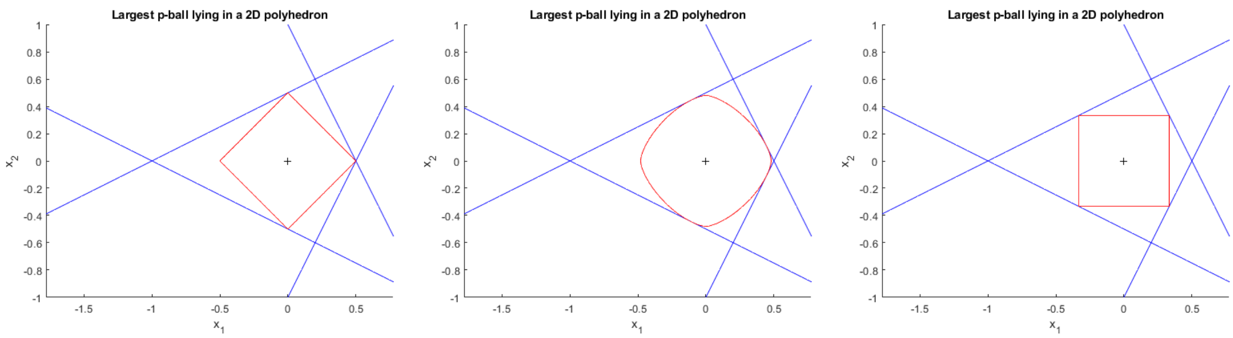

Figure 3: Left: $p = 1$; Middle: $p = 1.5$; Right: $p = \infty$.

```
function [x_c, r] = ChebyCenterQnorm(A, b, p)
% each column of A corresponds to a line
% --- input ---
% A: a matrix;
% b: a vector -- the RHS
% A(:,k) ... b(k)
% p: for p-norm
% --- output ---
```

```
% x_c: the location of the center
% r: the radius


n = size(A,2);
B = (A');
Babs = abs(B);
q = 1/(1-(1/p));

if p == 1
    len = max(Babs,[],2);
else
    len = (sum(Babs.^q ,2)).^(1/q);
end


% Create and solve the model
cvx_begin
    variable r(1)
    variable x_c(2)
    maximize ( r )
    B*x_c + r* len ≤ b;
cvx_end

% Generate the figure
x = linspace(-2,2);
%theta = 0:pi/100:2*pi;

hold on

for i = 1:n
    plot( x, -x*A(1,i)./A(2,i) + b(i)./A(2,i),'b-');
end
%plot( x_c(1) + r*cos(theta), x_c(2) + r*sin(theta), 'r');

up = r*(-1:0.01:1);
xp = up + x_c(1);
lxp = length(xp);
if p == inf
    xp = [-r *ones(1,lxp)+x_c(1),xp, r*ones(1,lxp) + x_c(1)];
    yup = [r*linspace(0,1,lxp), r*ones(1,lxp),  r-r*linspace(0,1,lxp)];
else
    yup = (r^p - abs(up).^p).^(1/p);
end
plot( xp, x_c(2) - yup, 'r');
plot( xp, x_c(2) + yup, 'r');

plot(x_c(1),x_c(2),'k+')
xlabel('x_1')
ylabel('x_2')
title('Largest p-ball lying in a 2D polyhedron');
axis([-1 1 -1 1])
axis equal
```

The main function is

```
close all;
clear variables;

A = [2,2,-1,-1; 1,-1,2,-2];
b = ones(4,1);

p= 1; %1.5 % Inf
[x_c, r] = ChebyCenterQnorm(A,b,p);
```

9. We cannot simply set p in the previous function when $p < 1$, because when formulating the LP, we use

the fact that

$$\sup \left\{ a_i^T u \,|\, \|u\|_p \leq r \right\} = r\|a_i\|_q.$$

This duality relation between p and q norms is based on Holder's inequality for $p \geq 1$, and is no longer true when $p < 1$. So the preceding problem formulation doesn't hold for $p < 1$.

If we try it, say $p = 0.5$, we will get the following plot.

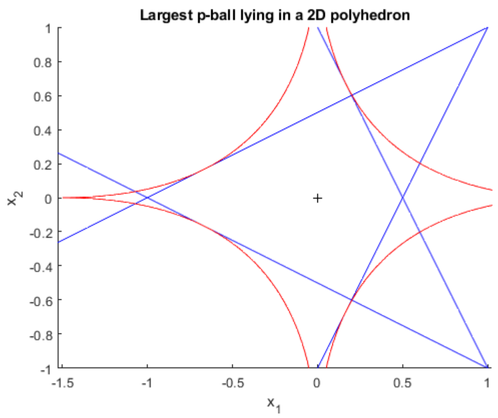

Figure 4: $p = 0.5$

10. [2] The polyhedron is a convex set. If a p-unit ball is in the polyhedron, then its convex hull must also lies in the polyhedron.

The key observation is that the convex hull of a p-unit ball is the 1-norm unit ball, at least in this $\mathbb{R}^2$ case. Since the 1-norm unit ball is convex, we only need to prove that it is a subset of any convex set containing the p-unit ball B. By symmetry, we only need to prove that the region where $x \geq 0, y \geq 0, x + y \leq 1$ is in the convex hull of B.

For any (x, y) such that $x \geq 0, y \geq 0, x + y \leq 1$, if $y < 1$, then the line going across (x, y) and $(0, 1)$ will hit the x-axis at $D = (x/(1 - y), 0)$. Since $0 \leq x/(1 - y) \leq 1$, $D \in B$. Since $(0, 1) \in B$, we have shown that such (x, y) is a convex combination of two points in B and thus, $(x, y) \in \mathbf{conv}B$.

Thus, the 1-norm unit ball is the convex hull of the p-unit ball when $p < 1$. So the polyhedron contains a radius r p-ball if and only if it contains a radius r 1-norm ball with the same center.

Therefore, when $p < 1$, we simply consider the case for $p = 1$ but draw the p-ball instead of 1-norm ball for the plot.

Here is an example with $p = 0.5$.

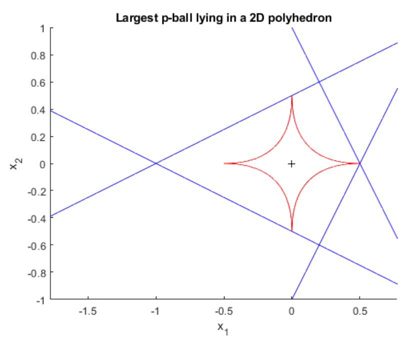

Figure 5: $p = 0.5$

```
function [x_c, r] = ChebyCenterQnormGeneral(A, b, p)
% p can be smaller than 1
% each column of A corresponds to a line
```

[2]Thanks for the discussion between Sonia and Prof. Overton on campuswire.

```
% --- input ---
% A: a matrix;
% b: a vector -- the RHS
% A(:,k) ... b(k)
% p: for p-norm
% --- output ---
% x_c: the location of the center
% r: the radius

n = size(A,2);
B = (A');
Babs = abs(B);
q = 1/(1-(1/p));

if p ≤ 1
    len = max(Babs,[],2);
else
    len = (sum(Babs.^q ,2)).^(1/q);
end


% Create and solve the model
cvx_begin
    variable r(1)
    variable x_c(2)
    maximize ( r )
    B*x_c + r* len ≤ b;
cvx_end

% Generate the figure
x = linspace(-2,2);
%theta = 0:pi/100:2*pi;

hold on

for i = 1:n
    plot( x, -x*A(1,i)./A(2,i) + b(i)./A(2,i),'b-');
end
%plot( x_c(1) + r*cos(theta), x_c(2) + r*sin(theta), 'r');

up = r*(-1:0.01:1);
xp = up + x_c(1);
lxp = length(xp);
if p == inf
    xp = [-r *ones(1,lxp)+x_c(1),xp, r*ones(1,lxp) + x_c(1)];
    yup = [r*linspace(0,1), r*ones(1,lxp),  r-r*linspace(0,1)];
else
    yup = (r^p - abs(up).^p).^(1/p);
end
plot( xp, x_c(2) - yup, 'r');
plot( xp, x_c(2) + yup, 'r');

plot(x_c(1),x_c(2),'k+')
xlabel('x_1')
ylabel('x_2')
title('Largest p-ball lying in a 2D polyhedron');
axis([-1 1 -1 1])
axis equal
```

The main function is

```
close all;
clear variables;


A = [2,2,-1,-1; 1,-1,2,-2];
```

```
b = ones(4,1);
p= 0.5;
[x_c, r] = ChebyCenterQnormGeneral(A,b,p);
```