CONVEX OPTIMIZATION HOMEWORK 4

TOMMENIX YU
ID: 12370130
STAT 31015
DUE WED FEB 1, 2023, 3PM

Exercise 1.

Proof.

(a): (discussed in office hour)

We know that the epigraph is convex, so we can find a supporting hyperplane at each $(x, f(x))$ since it is on the boundary of the epigraph. Now we can use the supporting hyperplane theorem to get that there exists $u_0 = (u', v)$ such that, for fixed x

$$u_0^T(y, f(y)) \geq u_0^T(x, f(x))$$

for all y since $(y, f(y))$ is a point in the epigraph.

Since can pick any point in the epigraph so we pick $(x, f(x) + 1)$. So we have

$$u_0^T(x, f(x) + 1) \geq u_0^T(x, f(x)) \Rightarrow v \cdot 1 \geq 0.$$

So v is non-negative. If $v = 0$, then we can pick $(x + u, |f(x + u)| + |f(x)| + 1)$, a point in the epigraph and get

$$u_0^T(x+u, |f(x+u)|+|f(x)|+1) \leq u_0^T(x+u, f(x+u)) \Rightarrow v(|f(x+u)|-f(x+u)+|f(x)|+1) \leq -||u||^2$$

which, since $u_0 \neq 0$, cannot be true since then it would be an instance of $0 < 0$.

If $v > 0$, then we we scale and define

$$u_1 := -\frac{1}{v}u_0 := (u, -1)$$

and thus since the sign flips when multiplied with a negative number

$$u_1^T(y, f(y)) \leq u_1^T(x, f(x)).$$

Now we have

$$u^T y - f(y) \leq y^T x - f(x) \Rightarrow f(y) \geq f(x) + u^T(y - x)$$

(b):

Let $u, v \in \partial f$, then we need to show that $\theta u + (1 - \theta)v \in \partial f$. Now we have for any y

$$\theta f(y) \geq \theta f(x) + \theta u^T(y - x)$$

and

$$(1-\theta)f(y) \geq (1-\theta)f(x) + (1-\theta)u^T(y-x)$$

which if we add them together we get that

$$f(y) \geq f(x) + (\theta u + (1-\theta)v)^T(y-x)$$

which means that ∂f is convex since $\theta u + (1-\theta)v \in \partial f$ for all $u, v \in \partial f$.

(c):

If $x = 0$ then $f(y) - f(x) \geq u(y-x)$ implies

$$|y| \geq uy$$

thus $u \in [-1, 1]$.

If $x > 0$ then we can pick $y > x$ with $y - x \leq \varepsilon$ to get that $u \leq 1$. Yet on the other hand we get when $y < x, x - y \leq \varepsilon$ and from the inequality we have $u \geq 1$. Thus $u = 1$ when $x > 0$.

When $x < 0$ by a similar argument we get that $u = -1$.

(Or we can say that u is the slope of the supporting plane, so result follows.)

$\square$

Exercise 2.

Proof.

$$\phi(\varepsilon) = \frac{f(x_0+\varepsilon) - f(x_0)}{\varepsilon}.$$

(a):

We compute the hessian of the function.

$$\frac{\partial f}{\partial x_i} = \lambda_i e^{-x_i}(1-e^{-x_i})^{-1} f$$

$$\frac{\partial^2 f}{\partial x_i^2} = -\frac{\lambda_i e^{-x_i}}{(1-e^{-x_i})^2} f + \frac{\lambda_i^2 e^{-2x_i}}{(1-e^{-x_i})^2} f$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \lambda_i e^{-x_i}(1-e^{-x_i})^{-1} \lambda_j e^{-x_j}(1-e^{-x_j})^{-1} f$$

And if we let $y_i = \dfrac{\lambda_i e^{-x_i}}{1-e^{-x_i}}$ then $y_i f = f_{\partial i}$. Again, if we let

$$S = \operatorname{diag}\left(\frac{\lambda_1 e^{-x_1}}{(1-e^{-x_1})^2}, \dots, \frac{\lambda_n e^{-x_n}}{(1-e^{-x_n})^2}\right)$$

then we have

$$\nabla^2_{xx} f = (yy^T - S) f$$

but since $f > 0$ and thus we need to prove that $yy^T - S$ is negative semi-definite, which by Schur decomposition we have

$$yy^T - S \preceq 0 \Rightarrow A := \begin{pmatrix} S & y \\ y^T & 1 \end{pmatrix} \preceq 0$$

i.e. the matrix below is negative. So we need to prove that for any u, $u^T A u = 0$. But we note that we can let

$$v := \left(u_1 \cdot (1-e^{-x_1}), \dots, u_n \cdot (1-e^{-x_n})\right)^T$$

which means that

$$u^T A u = 0 \iff v^T B v = 0$$

for

$$B := \begin{pmatrix} \lambda_1 e^{-x_1} & \dots & & & \lambda_1 e^{-x_1} \\ & \lambda_2 e^{-x_2} & & & \lambda_1 e^{-x_2} \\ \vdots & & \ddots & & \vdots \\ & & & \lambda_n e^{-x_n} & \lambda_1 e^{-x_2} \\ \lambda_1 e^{-x_1} & \lambda_2 e^{-x_2} & \dots & & 1 \end{pmatrix}$$

just by simple element wise distributing the denominators.

So to prove that f is concave, we need to show that the matrix A is negative semi definite, for which it suffices to show that B is negative semi definite since we can find a corresponding v for each u in the test for A.

But note that B is diagonally dominant because for every row except the only non-zero term except the diagonal is the same as the diagonal, and for the last row, we have that the domain is

$$\sum_{i=1}^{n} \lambda_i e^{-x_i} \leq 1.$$

Thus the Gershgorin circle theorem we know that all eigenvalues of B is with in the unions of balls (intervals) in which all the points are non negative, i.e. B is positive semi definite.

Thus we've shown that f is concave.

(b):

Taking the derivative we get

$$f' = \nabla g(x) \exp(-g(x))$$

and

$$\begin{aligned} f'' &= \nabla^2_{xx} g(x) \exp(-g(x)) - (\nabla g(x))^2 \exp(-g(x)) \\ &= \det \begin{pmatrix} \nabla^2_{xx} g(x) & \nabla g(x) \\ \nabla g(x)^T & 1 \end{pmatrix} \exp(-g(x)) \geq 0. \end{aligned}$$

Hence we are done.

$\square$

Exercise 3.

Proof.

We will use the property that a solution x is optimal $\iff$ it is feasible and

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

for all y feasible.

Thus, we first compute the derivative of $||Ax-b||^2$, for which we compute the gradient:

$$\nabla(Ax-b)^T(Ax-b) = 2A^T(Ax-b).$$

Yet since $l_i - x_i \leq (y-x)_i \leq u_i - x_i$ and we need

$$2(Ax-b)^T A(y-x) \geq 0$$

for all feasible y. Note that if for all i, $l_i - x_i \geq 0$ then we only need the gradient to be $\succeq 0$, and if the inequality is flipped we can then flip the requirement on $2(Ax-b)A^T$.

If for some i $l_i - x_i < 0$ then we can let $y = x - \epsilon e_i$ then we'll see that the dot product is negative. A similar argument applies to when $x_i < u_i$ for some i: we can always find a y with only one coordinate slightly larger or smaller than x. Thus the result is:

$$\begin{cases} A^T(Ax-b) = 0 & x \neq l and x \neq u \\ A^T(Ax-b) \preceq 0 & x = l \\ A^T(Ax-b) \succeq 0 & x = u \end{cases}$$

where $l = (l_1, \ldots, l_n)^T$ and $u = (u_1, \ldots, u_n)^T$.

$\square$

Exercise 4.

Proof.

(a):

For any given x_0, we know that the set $S(x_0)$ is compact. Thus since the function $f_0(x) + c_n||Ax - b||^2$ is continuous, it reaches its infimum in $S(x_0)$.

But for any point y outside the set $S(x_0)$, we know that $f(y) > f(x_0)$ and $||Ay - b||^2 \geq 0$, which in turn means that

$$f_0(x_0) + c_n||Ax_0 - b||^2 \leq f_0(y) + c_n||Ay - b||^2.$$

Thus, the global minimum is in the set $S(x_0)$, whose minimum is attained due to compactness.

(b):

Let's fix x_0, then we know from argument above that any point outside of $S(x_0)$ is not a solution of $\mathcal{P}_n$, for all n. But this tell us that x_n is a sequence in a compact set, which means that it has a converging subsequence, i.e. an accumulation point.

(c):

Fix x_0, then for any $x \in S(x_0)$, since $f(x)$ is fixed, if $||Ax - b||^2 = c > 0$, then there exists N such that for all $n \geq N$, $c_n||Ax - b||^2 \geq f(x_0) - f(x)$, which means x is not a solution of the problem $\mathcal{P}_n$. Moreover, since the function is continuous, thus x cannot be an accumulation point of x_n, since for all the same N above there exists a δ such that for any δ ball around x the point can be also eliminated. Thus there's at most N points in that ball, so x is not an accumulation point.

Hence, we've proven that x with $||Ax - b||^2 > 0$ is not an accumulation point, which means that any accumulation point is feasible.

(d):

Now we show that if there is an accumulation point with $||Ax - b||^2 = 0$, it cannot be non-optimal for f_0.

Suppose y satisfies $||Ax - b||^2 = 0$ is an accumulation point, then y is the minimal element of at least infinite problems $\mathcal{P}_n$. Yet for each $\mathcal{P}_n$ where y is a solution for, we know that the minimum point x of the original question has that

$$f_0(x) + c_n||Ax - b||^2 \leq f_0(y) + c_n||Ay - b||^2$$

where the equality is attained iff $x = y$ since the function is convex. But if the inequality is strict then y is not a solution, so the equality holds, i.e. y is an optimal point for the original point.

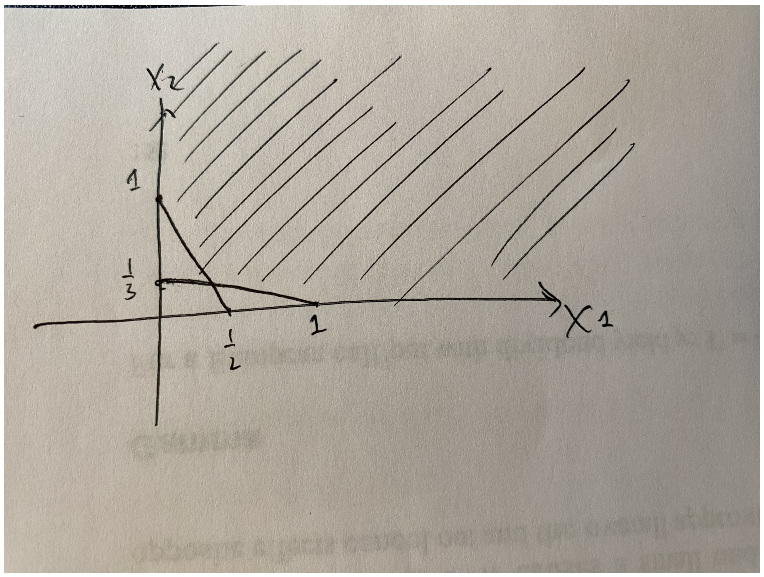
□

Exercise 5.

Proof.

(a):

The graph is:



c: Optimal set is the most left ward part of the domain, i.e.

$$\{(0, x_2)|x_2 \geq 1\}$$

and the optimal value is 0.

d: The optimal set is where the smallest box centered at the origin intersects the domain: $\left\{\left(\frac{1}{3}, \frac{1}{3}\right)\right\}$ and the optimal value is $\frac{1}{3}$.

e: The optimal set of $f_0 = x_1^2 + 9x_2^2$ is where the smallest ellipse $\frac{x^2}{9} + y^2 = c$ intersects the domain. Thus it must be either on the boundary segment $\left[(0, 1), \left(\frac{2}{5}, \frac{1}{5}\right)\right]$ or the segment $\left[\left(\frac{2}{5}, \frac{1}{5}\right), (1, 0)\right]$.

On the first line segment $x_2 = 1 - 2x_1$ and thus

$$f_0(x_1, x_2) = 37x^2 - 36x + 9 = \left(\sqrt{37}x - \frac{18}{\sqrt{37}}\right) + \frac{9}{37}$$

hence the (possible) optimal set is $\left\{\left(\frac{18}{37}, \frac{1}{37}\right)\right\}$ and the (possible) optimal value is $\frac{9}{37}$.

On the other line segment $x_1 = 1 - 3x_2$ and thus

$$f_0(x_1, x_2) = 18x_2^2 - 6x_2 + 1 = 6\left(\sqrt{3}x_2 - \frac{1}{2\sqrt{3}}\right) + \frac{1}{2}$$

hence the (possible) optimal set is $\left\{\left(\frac{1}{2}, \frac{1}{6}\right)\right\}$ and the (possible) optimal value is $\frac{1}{2}$.

But note that $\left(\frac{1}{2}, \frac{1}{6}\right)$ is on the boundary and $\left(\frac{18}{37}, \frac{1}{37}\right)$ is outside, so the real optimal point is $\left\{\left(\frac{1}{2}, \frac{1}{6}\right)\right\}$ and the real optimal value is $\frac{1}{2}$.

(b): The results are:

```
-------------------------------------------------

-------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +8.45293e-10
```

```
------------------------------------------------------------

------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +0.333333
```

```
 termination code       =  0
 DIMACS: 4.6e-14  0.0e+00  3.9e-12  0.0e+00  1.0e-08  1.0e-08
-------------------------------------------------------------------

------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +0.5
```

So the result is the same as computed.

(c): The code for the third problem is:

```

cvx_begin
variable x
variable y
minimize(x^2+9*y^2)
subject to
2*x+y >= 1;
x+3*y >= 1;
x >= 0;
y >= 0;
cvx_end

```

□