CONVEX OPTIMIZATION HOMEWORK 6

TOMMENIX YU
ID: 12370130
STAT 31015
DUE WED FEB 22, 2023, 3PM

Exercise 1.

Proof.

(1)

If the Slater's condition is not satisfied at point x^*, then for some i we have $f_i(x^*) = 0$, hence even if we take the tangent line (or linearization) at point x^*, we will still get $l_i(x^*) = 0$ since $f_i(x^*) = l_i(x^*)$. Thus, the Slater's condition is not satisfied for the linearized version.

(2)

If the Slater's condition is satisfied for the linearized version, we will have that the Slater's condition is satisfied for the original problem at x^*. Assume that it is satisfied, then the constraints of the linearized version can be rewritten into

$$l_i(x) < 0 \iff a_i \cdot x - b_i < 0 \iff Ax \prec b$$

which by example 2.21 is infeasible iff

$$\lambda \neq 0; \quad \lambda \succeq 0; \quad A^T\lambda = 0; \quad \lambda^T b \leq 0$$

is satisfied for some $\lambda \in \mathbb{R}^m$.

But note that the KKT condition for the linearized system with strict condition is

$$Ax < b; \quad \lambda \succeq 0; \quad \lambda_i(a_i \cdot x - b_i) = 0; \quad \nabla f_0(x) + A^T\lambda = 0$$

and so for the solution x^*, if the KKT conditions are satisfied for some $\lambda \neq 0$, then $\nabla f_0(x^*) + A^T\lambda = 0$ implies $A^T\lambda = 0$; for the last condition we know

$$\sum_{i=1}^{m} \lambda_i(a_i \cdot x^* - b_i) = \lambda^T A x^* - \lambda^T b = 0$$

so

$$\lambda^T b = \lambda^T A x^* = \left((x^*)^T (A^T\lambda)^T\right)^T = 0 \leq 0$$

so the conditions from 2.21 is satisfied, thus $Ax \prec b$ is infeasible for all x, hence not for x^*. So there is some i with $a_i x - b_i = 0$ and note that we can just add any multiple of e_i to λ to get still a λ satisfying the KKT conditions above (using infeasible we get $A^T\lambda = 0$ for each one), so we can construct a sequence of $\lambda \to \infty$. So the dual problem is unbounded. □

Exercise 2.

Proof.

Assume that the matrix is non-singular, then there exists some $(x, y) \neq 0$ such that it is in the kernel of the matrix, that is

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

If $x = 0$ then we know $A^T y = 0$ since A^T is full column rank, thus $y = 0$ contradiction. So $x \neq 0$. But then we have

$$Ax = 0, x \neq 0, Hx + A^T y = 0$$

where multiplying on left with x^T to the last equation gives

$$x^T Hx + (y^T Ax)^T = 0 \Rightarrow 0 < x^T Hx = 0$$

contradiction!

So there's no such non-zero vector in it's kernel, thus it's non-singular.

$\square$

Exercise 3.

Proof.

Taking the derivative we get

$$\nabla^2 g(x) = \frac{1}{c^2} \exp\left(-f(x)/c\right) \left(-c\nabla^2 f(x) + f(x) f(x)^T\right).$$

The condition $\lambda^2 \geq c$ explicitly written is

$$0 < \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \leq c$$

which implies, by Schur decomposition, along with the condition that $\nabla^2 f(x) \succ 0$ that

$$\begin{pmatrix} \nabla^2 f(x) & \nabla f(x) \\ \nabla f(x)^T & c \end{pmatrix} \geq 0.$$

But this implies

$$0 \leq \begin{pmatrix} u^T & a \end{pmatrix} \begin{pmatrix} \nabla^2 f(x) & \nabla f(x) \\ \nabla f(x)^T & c \end{pmatrix} \begin{pmatrix} u \\ a \end{pmatrix} = u^T \nabla^2 f(x) u + 2(u^T \nabla f(x)) a + a^2 c$$

hence we use $a = \dfrac{-u^T \nabla f}{c}$ to get

$$u^T \nabla^2 f(x) u - 2(u^T \nabla f(x)) \frac{u^T \nabla f}{c} + \frac{(u^T \nabla f)^2}{c^2} c = u^T \left(\nabla^2 f(x) - \frac{1}{c} \nabla f(x) \nabla f(x)^T\right) u \geq 0$$

where since $c > 0$ again we have

$$\left(c\nabla^2 f(x) - \nabla f(x) \nabla f(x)^T\right) \geq 0$$

which means

$$\nabla^2 g(x) = -\frac{1}{c^2} \exp\left(-f(x)/c\right) \left(c\nabla^2 f(x) - f(x) f(x)^T\right) \leq 0.$$

$\square$

Exercise 4.

Proof.

So I followed the instructions and wrote up the codes (see Github file, since it's a bit long for screen shot here, and a lot), with A **generated by "randn" code** (I notice that with "rand" the result is different for a fixed seed).

Moreover, in my files, I followed the instructions in the assignment requirement that

- $[x1, \lambda_1]$ are for Newton method (a) starting at x_0;
- $[x2, \lambda_2]$ are for infeasible Newton method (b) starting at $\hat{x}$, which is not dependent upon the input x_0;
- $[x3, \lambda_3]$ are for infeasible Newton method (b) starting at x_0.

As for the result, when I used $\hat{x}$ as the starting point, the two methods indeed coincide (only the first few elements of the solution point here):

```
ans =

    0.3222    0.2503    0.6165    0.4095    0.4439    0.6679    0.3168    0.1855    0.7105    0.3969

ans =

    0.3222    0.2503    0.6165    0.4095    0.4439    0.6679    0.3168    0.1855    0.7105    0.3969
```

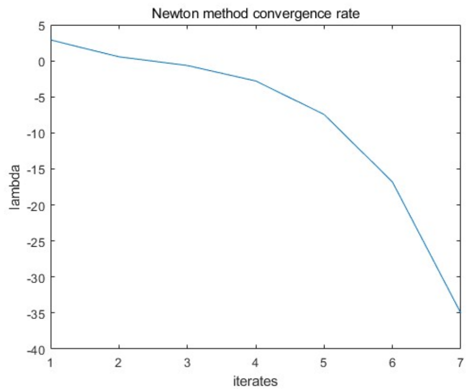
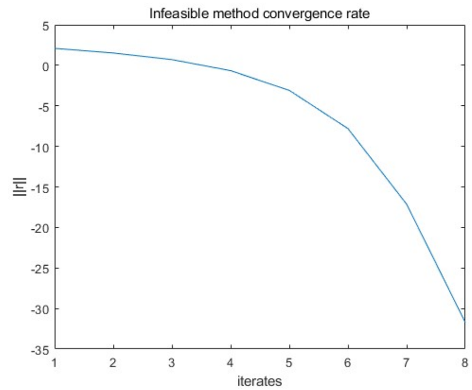
For starting point 1 it's the same:

```
>> text1

x =

   0.3222 - 0.0000i
   0.2503 + 0.0000i
   0.6165 + 0.0000i
   0.4095 + 0.0000i
   0.4439 - 0.0000i
   0.6679 - 0.0000i
   0.3168 - 0.0000i
   0.1855 - 0.0000i
   0.7105 - 0.0000i
   0.3969 + 0.0000i
   0.5167 + 0.0000i
   0.3155 - 0.0000i
```

And they are the same. Moreover, the convergence is indeed quadratic for both:

Newton	Infeasible
	

□