COURSEWORK 2

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

477 - Computational Optimisation

Coursework 2

Author:
Jiahao Lin (CID: 00837321)

Date: November 21, 2016

1 Part 1

1.1 Q.1

a) To show that $\mathbf{\Delta} x_k$ is descent direction at x_k, we need to show that:

$$\nabla f(x_k)^\top \mathbf{\Delta} x_k < 0 \tag{1}$$
$$\tag{2}$$

Since:

$$\nabla^2 f(x_k)\mathbf{\Delta} x_k = -\nabla f(x_k) \tag{3}$$
$$\tag{4}$$

We can get:

$$\nabla f(x_k)^\top \mathbf{\Delta} x_k = -\nabla f(x_k)^\top \frac{\nabla f(x_k)}{\nabla^2 f(x_k)} \tag{5}$$
$$= -\frac{\nabla f(x_k)^\top \nabla f(x_k)}{\nabla^2 f(x_k)} \tag{6}$$

We can see that this is less than zero since the numerator is positive definite and the same for denominator:

$$\nabla^2 f(x_k) \geq m\boldsymbol{I} \tag{7}$$

b) We can set tolerane to e^{-08} and say that:

$$|f(x_{k+1}) - f(x_k)| < tor \tag{8}$$
$$||\nabla f(x_k)||_2 < tor \tag{9}$$
$$||x_{k+1} - x_k||_2 < tor \tag{10}$$

This is to check that the First Order Necessary Condition is satisfied.

c) Yes, since the function is strongly convex, but the condition is that the initial point x_0 has to be close enough to the optimal point.

d) First say that:

$$x_{k+1} = x_k + t_k \mathbf{\Delta} x_k \tag{11}$$

Then

$$f(x_{k+1}) = f(x_k + t_k \mathbf{\Delta} x_k) \tag{12}$$

Now use Taylor expansion to expand the above function into second order:

$$f(x_k + t_k \mathbf{\Delta} x_k) \approx f(x_k) + t_k \langle \nabla f(x_k), \mathbf{\Delta} x_k \rangle + \frac{1}{2}\nabla^2 f(x_k)||t_k||_2^2||\mathbf{\Delta} x_k||_2^2 \tag{13}$$

$$\Delta x_k = -\frac{\nabla f(x_k)}{\nabla^2 f(x_k)} \tag{14}$$

$$\langle \nabla f(x_k), \Delta x_k \rangle = -\frac{||\nabla f(x_k)||_2^2}{\nabla^2 f(x_k)} \tag{15}$$

We need to show that:

$$f(x_k) + t_k\langle \nabla f(x_k), \Delta x_k\rangle + \frac{1}{2}\nabla^2 f(x_k)||t_k||_2^2||\Delta x_k||_2^2 \leq f(x_k) + \alpha t_k\langle \nabla f(x_k), \Delta x_k)\rangle \tag{16}$$

$$t_k\langle \nabla f(x_k), \Delta x_k)\rangle + \frac{1}{2}\nabla^2 f(x_k)||t_k||_2^2||\Delta x_k||_2^2 \leq \alpha t_k\langle \nabla f(x_k), \Delta x_k)\rangle \tag{17}$$

Now let's look at first term fisrt:

$$t_k\langle \nabla f(x_k), \Delta x_k)\rangle \tag{18}$$

Since we have the condition on t_k:

$$t \leq -\frac{\langle \nabla f(x_k), \Delta x_k)\rangle}{\boldsymbol{M}||\nabla x_k||_2^2} \tag{19}$$

$$0 > \langle \nabla f(x_k), \Delta x_k)\rangle t_k > -\langle \nabla f(x_k), \Delta x_k)\rangle^\top \frac{\langle \nabla f(x_k), \Delta x_k)\rangle}{\boldsymbol{M}||\nabla x_k||_2^2} \tag{20}$$

$$0 > \langle \nabla f(x_k), \Delta x_k)\rangle t_k > -\frac{||\nabla f(x_k)||^2||\Delta x_k)||^2}{\boldsymbol{M}||\nabla x_k||_2^2} \tag{21}$$

$$0 > \langle \nabla f(x_k), \Delta x_k)\rangle t_k > -\frac{||\nabla f(x_k)||^2}{\boldsymbol{M}} \tag{22}$$

Now let's take a look at second term:

$$\frac{1}{2}\nabla^2 f(x_k)||t_k||_2^2||\Delta x_k||_2^2 \tag{23}$$

Again substitute in condition for t:

$$t \leq -\frac{\langle \nabla f(x_k), \Delta x_k)\rangle}{\boldsymbol{M}||\nabla x_k||_2^2} \tag{24}$$

$$||t||_2^2 \leq \frac{||\nabla f(x_k)||^2||\Delta x_k)||^2}{\boldsymbol{M}^\top\boldsymbol{M}||\nabla x_k||_2^4} \tag{25}$$

$$||t||_2^2 \leq \frac{||\nabla f(x_k)||^2}{\boldsymbol{M}^\top\boldsymbol{M}||\nabla x_k||_2^2} \tag{26}$$

$$||t||_2^2||\Delta x_k||_2^2 \leq \frac{||\nabla f(x_k)||^2}{\boldsymbol{M}^\top\boldsymbol{M}} \tag{27}$$

$$\tag{28}$$

Since we have the upper bound for $\nabla^2 f(x_k)$:

$$\nabla^2 f(x_k) \leq \boldsymbol{MI} \tag{29}$$

$$||t||_2^2||\Delta x_k||_2^2\nabla^2 f(x_k) \leq \frac{\nabla^2 f(x_k)||\nabla f(x_k)||^2}{\boldsymbol{M}^\top\boldsymbol{M}} \tag{30}$$

$$\tag{31}$$

e)

2 Part 2

2.1 Q.2

a) KKT conditions:

$$min \quad z_1^2 + (x_2+1)^2 \tag{32}$$
$$g(x^*) = exp(x_1^*) - x_2^* \leq 0 \tag{33}$$
$$\mu^* \geq 0 \tag{34}$$
$$2x_1^* + \mu^* exp(x_1^*) = 0 \tag{35}$$
$$2(x_2^* + 1) + \mu^* * (-1) = 0 \tag{36}$$
$$\mu^*(exp(x_1^*) - x_2^*) = 0 \tag{37}$$

if $\mu^* = 0$:

$$2x_1^* = 0 \tag{38}$$
$$x_1^* = 0 \tag{39}$$
$$2(x_2^* + 1) = 0 \tag{40}$$
$$x_2^* = -1 \tag{41}$$
$$g(x^*) = exp(0) - (-1) = 2 > 0 \tag{42}$$
$$\tag{43}$$

This is contradicting our condition, so $\mu^* > 0$:

$$exp(x_1^*) - x_2^* = 0 \tag{44}$$
$$exp(x_1^*) = x_2^* \tag{45}$$
$$2(x_2^* + 1) + \mu^*(-1) = 0 \tag{46}$$
$$2(exp(x_1^*) + 1) = \mu^* \tag{47}$$
$$2x_1^* + \mu^* * exp(x_1^*) = 0 \tag{48}$$
$$2x_1^* + 2(exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{49}$$

Since $exp(x_1^*) > 0$:

$$2(exp(x_1^*) + 1) * exp(x_1^*) > 0 \tag{50}$$
$$\tag{51}$$

combine with

$$2x_1^* + 2(exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{52}$$

we got

$$2x_1^* < 0 \tag{53}$$
$$x_1^* < 0 \tag{54}$$
$$0 < exp(x_1^*) < 1 \tag{55}$$

$$1 < exp(x_1^*) + 1 < 2 \tag{56}$$
$$0 < (exp(x_1^*) + 1) * exp(x_1^*) < 2 \tag{57}$$
$$\tag{58}$$

combine with the above equation agian:

$$x_1^* + (exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{59}$$
$$-1 < x_1^* < 0 \tag{60}$$

b) KKT condition:

$$min \quad c^\top x + 8 \tag{61}$$
$$g(x^*) = \frac{1}{2}||x^*||^2 - 1 \leq 0 \tag{62}$$
$$\mu^* \geq 0 \tag{63}$$
$$c + \mu^* x^* = 0 \tag{64}$$
$$\mu^*(\frac{1}{2}||x^*||^2 - 1) = 0 \tag{65}$$

We can see that if $\mu^* = 0$, $c = 0$ as well, which contradicts our given condition $c \neq 0$, so we get $\mu^* > 0$

$$\frac{1}{2}||x^*||^2 - 1 = 0 \tag{66}$$
$$||x^*||^2 = 2 \tag{67}$$
$$||a\mathbf{1}||^2 = 2 \tag{68}$$
$$n|a|^2 = 2 \tag{69}$$
$$a = \pm\sqrt{\frac{2}{n}} \tag{70}$$
$$c^\top x + 8 = 4 \tag{71}$$
$$c^\top x = -4 \tag{72}$$
$$n(\boldsymbol{c}a) = -4 \tag{73}$$
$$\boldsymbol{c} = \frac{-4}{an} \tag{74}$$
$$c = \mp\frac{4}{\sqrt{2n}} \tag{75}$$

Please noted that here c and a have inverse signs.

c) KKT Condition For original problem:

$$min \quad f(x) \tag{76}$$
$$h(x^*) = 0 \tag{77}$$
$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0 \tag{78}$$

KKT Condition For In-Eq problem:

$$g(x^*) = \frac{1}{2}||h(x^*)||^2 \leq 0 \tag{79}$$
$$\mu^* \geq 0 \tag{80}$$
$$\nabla f(x^*) + \mu^* h(x^*) = 0 \tag{81}$$
$$\frac{1}{2}\mu^* ||h(x^*)||^2 = 0 \tag{82}$$

If $\mu^* = 0$:

$$\nabla f(x^*) = 0 h(x^*) may not be 0 \tag{83}$$

So $\mu^* > 0$:

$$||h(x^*)||^2 = 0 \nabla g(x^*) = ||h(x^*)|| = 0 h(x^*) = 0 \tag{84}$$

So $h(x^*)$ is not linearly independent, and x^* is not a regular point, so KKT theorem can not be applied on In-Eq problem.