COURSEWORK 2

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

477 - Computational Optimisation

Coursework 2

Author:
Jiahao Lin (CID: 00837321)

Date: November 23, 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 $1.0e^{-08}$ 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:

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

$$\Delta x_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k) \tag{14}$$

$$\langle \nabla f(x_k), \Delta x_k \rangle = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)^\top \nabla f(x_k) \tag{15}$$

We need to show that:

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

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

$$\frac{1}{2} \nabla^2 f(x_k) ||t_k||_2^2 ||\Delta x_k||_2^2 \leq (\alpha - 1) t_k \langle \nabla f(x_k), \Delta x_k \rangle \tag{18}$$

$$\frac{1}{2} \nabla^2 f(x_k) t_k ||\Delta x_k||_2^2 \leq (\alpha - 1) \langle \nabla f(x_k), \Delta x_k \rangle \tag{19}$$

Since we have condition on α:

$$0 < \alpha < 0.5 \tag{20}$$

$$-1 < \alpha - 1 < -0.5 \tag{21}$$

Since $\langle \nabla f(x_k), \Delta x_k) \rangle < 0$:

$$-\langle \nabla f(x_k), \Delta x_k) \rangle \geq (\alpha - 1) \langle \nabla f(x_k), \Delta x_k) \rangle \geq -0.5 \langle \nabla f(x_k), \Delta x_k \rangle \tag{22}$$

So we need to show that:

$$\frac{1}{2} \nabla^2 f(x_k) t_k ||\Delta x_k||_2^2 \leq -\langle \nabla f(x_k), \Delta x_k \rangle \tag{23}$$

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

$$t_k \leq -2 \frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{\nabla^2 f(x_k) ||\Delta x_k||_2^2} \tag{25}$$

Since we have condition on $\nabla^2 f(x_k)$:

$$\boldsymbol{mI} \preceq \nabla^2 f(x_k) \preceq \boldsymbol{MI} \tag{26}$$

$$\frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{\nabla^2 f(x_k) ||\Delta x_k||_2^2} \geq \frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{\boldsymbol{M} ||\Delta x_k||_2^2} \tag{27}$$

$$-\frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{\nabla^2 f(x_k) ||\Delta x_k||_2^2} \leq -\frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{\boldsymbol{M} ||\Delta x_k||_2^2} \tag{28}$$

$$\tag{29}$$

We got:

$$t_k \leq -2 \frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{M ||\Delta x_k||_2^2} \tag{30}$$

From the condition we know:

$$t_k \leq -\frac{\langle \nabla f(x_k), \Delta x_k) \rangle}{M ||\Delta x_k||_2^2} \tag{31}$$

Since t_k is positive, so we know the final equation we shown is true.

e) Assume t_0 is the initial step size, according to the condition given for Hessian of $f(x_k)$

$$\boldsymbol{mI} \preceq \nabla^2 f(x_k) \preceq \boldsymbol{MI} \tag{32}$$
$$\tag{33}$$

We can deduce:

$$\boldsymbol{mI} \preceq \nabla^2 f(x_k) \preceq \boldsymbol{MI} \tag{34}$$
$$\boldsymbol{mI} \preceq \frac{\nabla f(x_k)}{x_k} \preceq \boldsymbol{MI} \tag{35}$$
$$\boldsymbol{m} \leq \frac{\langle \nabla f(x_k), x_k \rangle}{||x_k||^2} \leq \boldsymbol{M} \tag{36}$$
$$\frac{\boldsymbol{m}}{\boldsymbol{M}} \leq \frac{\langle \nabla f(x_k), x_k \rangle}{\boldsymbol{M}||x_k||^2} \leq 1 \tag{37}$$
$$\frac{\boldsymbol{m}}{\boldsymbol{M}} \geq -\frac{\langle \nabla f(x_k), x_k \rangle}{\boldsymbol{M}||x_k||^2} \geq 1 \tag{38}$$
$$\tag{39}$$

According to the upper bound for t_k when backtracking stop, we got that when

$$t_k \leq \frac{\boldsymbol{m}}{\boldsymbol{M}} \tag{40}$$

backtracking stops, which is equivalent to

$$\beta^n t_0 \leq \frac{\boldsymbol{m}}{\boldsymbol{M}} \tag{41}$$

which the n_0 is maximum number of iteration for backtracking, and it could be expressed as:

$$n_0 = min \quad n \quad st. \quad n \leq \log_\beta \frac{\boldsymbol{m}}{\boldsymbol{M} t_0} = \log_\beta \boldsymbol{m} - \log_\beta \boldsymbol{M} - \log_\beta t_0 \tag{42}$$

2 Part 2

2.1 Q.2

a) KKT conditions:

$$min \quad z_1^2 + (x_2+1)^2 \tag{43}$$
$$g(x^*) = exp(x_1^*) - x_2^* \leq 0 \tag{44}$$
$$\mu^* \geq 0 \tag{45}$$
$$2x_1^* + \mu^* exp(x_1^*) = 0 \tag{46}$$
$$2(x_2^*+1) + \mu^* * (-1) = 0 \tag{47}$$
$$\mu^*(exp(x_1^*) - x_2^*) = 0 \tag{48}$$

if $\mu^* = 0$:

$$
\begin{aligned}
2x_1^* &= 0 && (49)\\
x_1^* &= 0 && (50)\\
2(x_2^*+1) &= 0 && (51)\\
x_2^* &= -1 && (52)\\
g(x^*) = exp(0)-(-1) &= 2 > 0 && (53)\\
& && (54)
\end{aligned}
$$

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

$$
\begin{aligned}
exp(x_1^*) - x_2^* &= 0 && (55)\\
exp(x_1^*) &= x_2^* && (56)\\
2(x_2^*+1) + \mu^*(-1) &= 0 && (57)\\
2(exp(x_1^*)+1) &= \mu^* && (58)\\
2x_1^* + \mu^* * exp(x_1^*) &= 0 && (59)\\
2x_1^* + 2(exp(x_1^*)+1) * exp(x_1^*) &= 0 && (60)
\end{aligned}
$$

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

$$
\begin{aligned}
2(exp(x_1^*)+1) * exp(x_1^*) > 0 && (61)\\
&& (62)
\end{aligned}
$$

combine with

$$
2x_1^* + 2(exp(x_1^*)+1) * exp(x_1^*) = 0 \tag{63}
$$

we got

$$
\begin{aligned}
2x_1^* &< 0 && (64)\\
x_1^* &< 0 && (65)\\
0 < exp(x_1^*) &< 1 && (66)\\
1 < exp(x_1^*)+1 &< 2 && (67)\\
0 < (exp(x_1^*)+1) * exp(x_1^*) &< 2 && (68)\\
& && (69)
\end{aligned}
$$

combine with the above equation agian:

$$
\begin{aligned}
x_1^* + (exp(x_1^*)+1) * exp(x_1^*) &= 0 && (70)\\
-2 < x_1^* &< 0 && (71)
\end{aligned}
$$

b) KKT condition:

$$
min \quad c^\top x + 8 \tag{72}
$$

$$g(x^*) = \frac{1}{2}\|x^*\|^2 - 1 \leq 0 \tag{73}$$
$$\mu^* \geq 0 \tag{74}$$
$$c + \mu^* x^* = 0 \tag{75}$$
$$\mu^*(\frac{1}{2}\|x^*\|^2 - 1) = 0 \tag{76}$$

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

$$\frac{1}{2}\|x^*\|^2 - 1 = 0 \tag{77}$$
$$\|x^*\|^2 = 2 \tag{78}$$
$$\|a\mathbf{1}\|^2 = 2 \tag{79}$$
$$n|a|^2 = 2 \tag{80}$$
$$a = \pm\sqrt{\frac{2}{n}} \tag{81}$$
$$c^\top x + 8 = 4 \tag{82}$$
$$c^\top x = -4 \tag{83}$$
$$n(\boldsymbol{c}a) = -4 \tag{84}$$
$$\boldsymbol{c} = \frac{-4}{an} \tag{85}$$
$$c = \mp\frac{4}{\sqrt{2n}} \tag{86}$$

Please noted that here c and a have inverse signs.

c) KKT Condition For original problem:

$$min \quad f(x) \tag{87}$$
$$h(x^*) = 0 \tag{88}$$
$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0 \tag{89}$$

KKT Condition For In-Eq problem:

$$g(x^*) = \frac{1}{2}\|h(x^*)\|^2 \leq 0 \tag{90}$$
$$\mu^* \geq 0 \tag{91}$$
$$\nabla f(x^*) + \mu^* h(x^*) = 0 \tag{92}$$
$$\frac{1}{2}\mu^*\|h(x^*)\|^2 = 0 \tag{93}$$

If $\mu^* = 0$:

$$\nabla f(x^*) = 0 \tag{94}$$

$$h(x^*) may not be 0 \tag{95}$$

So $\mu^* > 0$:

$$||h(x^*)||^2 = 0 \tag{96}$$
$$\nabla g(x^*) = ||h(x^*)|| = 0 \tag{97}$$
$$h(x^*) = 0 \tag{98}$$

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.