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 Δx_k is descent direction at x_k, we need to show that:

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

$$\tag{2}$$

Since:

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

$$\tag{4}$$

We can get:

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

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

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

$$\nabla^2 f(x_k) \succeq m\mathbf{I} \tag{7}$$

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:

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

Then

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

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

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

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

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

We need to show that:

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

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

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

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

Since we have condition on α:

$$0 < \alpha < 0.5 \tag{20}$$
$$-1 < \alpha - 1 < -0.5 \tag{21}$$

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

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

So we need to show that:

$$\frac{1}{2}\nabla^2 f(x_k)t_k\|\Delta x_k\|_2^2 \le -0.5\langle \nabla f(x_k), \Delta x_k \rangle \tag{23}$$

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

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

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

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

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

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

$$\tag{29}$$

We got:

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

Which is the given condition on t_k

e) Assume $t_0 = 1$, maximum number of backtracking step is max n that:

$$1 * \beta^n \leq -\frac{\langle \nabla f(x_n), \Delta x_n \rangle}{M\|\Delta x_n\|_2^2} \tag{31}$$

$$\beta^n \leq -\frac{\langle \nabla f(x_n), \Delta x_n \rangle}{M\|\Delta x_n\|_2^2} \tag{32}$$

$$n \leq \log_\beta -\frac{\langle \nabla f(x_n), \Delta x_n \rangle}{M\|\Delta x_n\|_2^2} \tag{33}$$

Maximum n satisfy the above condition is upper bound for iteration of backtracking steps.

2   Part 2

2.1   Q.2

a) KKT conditions:

$$min \quad z_1^2 + (x_2 + 1)^2 \tag{34}$$
$$g(x^*) = exp(x_1^*) - x_2^* \leq 0 \tag{35}$$
$$\mu^* \geq 0 \tag{36}$$
$$2x_1^* + \mu^* exp(x_1^*) = 0 \tag{37}$$
$$2(x_2^* + 1) + \mu^* * (-1) = 0 \tag{38}$$
$$\mu^*(exp(x_1^*) - x_2^*) = 0 \tag{39}$$

if $\mu^* = 0$:

$$2x_1^* = 0 \tag{40}$$
$$x_1^* = 0 \tag{41}$$
$$2(x_2^* + 1) = 0 \tag{42}$$
$$x_2^* = -1 \tag{43}$$
$$g(x^*) = exp(0) - (-1) = 2 > 0 \tag{44}$$
$$\tag{45}$$

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

$$exp(x_1^*) - x_2^* = 0 \tag{46}$$
$$exp(x_1^*) = x_2^* \tag{47}$$
$$2(x_2^* + 1) + \mu^*(-1) = 0 \tag{48}$$
$$2(exp(x_1^*) + 1) = \mu^* \tag{49}$$
$$2x_1^* + \mu^* * exp(x_1^*) = 0 \tag{50}$$
$$2x_1^* + 2(exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{51}$$

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

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

$$\tag{53}$$

combine with

$$2x_1^* + 2(exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{54}$$

we got

$$2x_1^* < 0 \tag{55}$$

$$x_1^* < 0 \tag{56}$$

$$0 < exp(x_1^*) < 1 \tag{57}$$

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

$$0 < (exp(x_1^*) + 1) * exp(x_1^*) < 2 \tag{59}$$

$$\tag{60}$$

combine with the above equation agian:

$$x_1^* + (exp(x_1^*) + 1) * exp(x_1^*) = 0 \tag{61}$$

$$-1 < x_1^* < 0 \tag{62}$$

b) KKT condition:

$$min \quad c^\top x + 8 \tag{63}$$

$$g(x^*) = \frac{1}{2}\|x^*\|^2 - 1 \le 0 \tag{64}$$

$$\mu^* \ge 0 \tag{65}$$

$$c + \mu^* x^* = 0 \tag{66}$$

$$\mu^*(\frac{1}{2}\|x^*\|^2 - 1) = 0 \tag{67}$$

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

$$\frac{1}{2}\|x^*\|^2 - 1 = 0 \tag{68}$$

$$\|x^*\|^2 = 2 \tag{69}$$

$$\|a\mathbf{1}\|^2 = 2 \tag{70}$$

$$n|a|^2 = 2 \tag{71}$$

$$a = \pm\sqrt{\frac{2}{n}} \tag{72}$$

$$c^\top x + 8 = 4 \tag{73}$$

$$c^\top x = -4 \tag{74}$$

$$n(\boldsymbol{ca}) = -4 \tag{75}$$

$$\boldsymbol{c} = \frac{-4}{an} \tag{76}$$

$$c = \mp\frac{4}{\sqrt{2n}} \tag{77}$$

Please noted that here c and a have inverse signs.

c) KKT Condition For original problem:

$$min \quad f(x) \tag{78}$$

$$h(x^*) = 0 \tag{79}$$

$$\nabla f(x^*) + \lambda^*\nabla h(x^*) = 0 \tag{80}$$

KKT Condition For In-Eq problem:

$$g(x^*) = \frac{1}{2}\|h(x^*)\|^2 \le 0 \tag{81}$$

$$\mu^* \ge 0 \tag{82}$$

$$\nabla f(x^*) + \mu^*h(x^*) = 0 \tag{83}$$

$$\frac{1}{2}\mu^*\|h(x^*)\|^2 = 0 \tag{84}$$

If $\mu^* = 0$:

$$\nabla f(x^*) = 0 \tag{85}$$

$$h(x^*)\,maynotbe0 \tag{86}$$

So $\mu^* > 0$:

$$\|h(x^*)\|^2 = 0 \tag{87}$$

$$\nabla g(x^*) = \|h(x^*)\| = 0 \tag{88}$$

$$h(x^*) = 0 \tag{89}$$

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.