

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:

$$\nabla f(x_k)^\top \Delta x_k < 0 \quad (1)$$

$$(2)$$

Since:

$$\nabla^2 f(x_k) \Delta x_k = -\nabla f(x_k) \quad (3)$$

$$(4)$$

We can get:

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

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

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

$$\nabla^2 f(x_k) \geq mI \quad (7)$$

b) We can set tolerane to $1.0e^{-08}$ and say that:

$$|f(x_{k+1}) - f(x_k)| < tor \quad (8)$$

$$\|\nabla f(x_k)\|_2 < tor \quad (9)$$

$$\|x_{k+1} - x_k\|_2 < tor \quad (10)$$

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:

$$x_{k+1} = x_k + t_k \Delta x_k \quad (11)$$

Then

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

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

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

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

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

We need to show that:

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

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

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

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

Since we have condition on α :

$$0 < \alpha < 0.5 \quad (20)$$

$$-1 < \alpha - 1 < -0.5 \quad (21)$$

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

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

So we need to show that:

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

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

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

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

$$mI \leq \nabla^2 f(x_k) \leq MI \quad (26)$$

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

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

$$(29)$$

We got:

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

From the condition we know:

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

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

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

$$1 * \beta^n \leq -\frac{\langle \nabla f(x_n), \Delta x_n \rangle}{M \|\Delta x_n\|_2^2} \quad (32)$$

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

$$n \leq \log_{\beta} -\frac{\langle \nabla f(x_n), \Delta x_n \rangle}{M \|\Delta x_n\|_2^2} \quad (34)$$

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

2 Part 2

2.1 Q.2

a) KKT conditions:

$$\min \quad z_1^2 + (x_2 + 1)^2 \quad (35)$$

$$g(x^*) = \exp(x_1^*) - x_2^* \leq 0 \quad (36)$$

$$\mu^* \geq 0 \quad (37)$$

$$2x_1^* + \mu^* \exp(x_1^*) = 0 \quad (38)$$

$$2(x_2^* + 1) + \mu^* * (-1) = 0 \quad (39)$$

$$\mu^* (\exp(x_1^*) - x_2^*) = 0 \quad (40)$$

if $\mu^* = 0$:

$$2x_1^* = 0 \quad (41)$$

$$x_1^* = 0 \quad (42)$$

$$2(x_2^* + 1) = 0 \quad (43)$$

$$x_2^* = -1 \quad (44)$$

$$g(x^*) = \exp(0) - (-1) = 2 > 0 \quad (45)$$

$$(46)$$

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

$$\exp(x_1^*) - x_2^* = 0 \quad (47)$$

$$\exp(x_1^*) = x_2^* \quad (48)$$

$$2(x_2^* + 1) + \mu^* (-1) = 0 \quad (49)$$

$$2(\exp(x_1^*) + 1) = \mu^* \quad (50)$$

$$2x_1^* + \mu^* \exp(x_1^*) = 0 \quad (51)$$

$$2x_1^* + 2(\exp(x_1^*) + 1) * \exp(x_1^*) = 0 \quad (52)$$

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

$$2(\exp(x_1^*) + 1) * \exp(x_1^*) > 0 \quad (53)$$

$$(54)$$

combine with

$$2x_1^* + 2(\exp(x_1^*) + 1) * \exp(x_1^*) = 0 \quad (55)$$

we got

$$2x_1^* < 0 \quad (56)$$

$$x_1^* < 0 \quad (57)$$

$$0 < \exp(x_1^*) < 1 \quad (58)$$

$$1 < \exp(x_1^*) + 1 < 2 \quad (59)$$

$$0 < (\exp(x_1^*) + 1) * \exp(x_1^*) < 2 \quad (60)$$

$$(61)$$

combine with the above equation again:

$$x_1^* + (\exp(x_1^*) + 1) * \exp(x_1^*) = 0 \quad (62)$$

$$-2 < x_1^* < 0 \quad (63)$$

b) KKT condition:

$$\min \quad c^\top x + 8 \quad (64)$$

$$g(x^*) = \frac{1}{2} \|x^*\|^2 - 1 \leq 0 \quad (65)$$

$$\mu^* \geq 0 \quad (66)$$

$$c + \mu^* x^* = 0 \quad (67)$$

$$\mu^* \left(\frac{1}{2} \|x^*\|^2 - 1 \right) = 0 \quad (68)$$

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

$$\frac{1}{2} \|x^*\|^2 - 1 = 0 \quad (69)$$

$$\|x^*\|^2 = 2 \quad (70)$$

$$\|a\mathbf{1}\|^2 = 2 \quad (71)$$

$$n|a|^2 = 2 \quad (72)$$

$$a = \pm \sqrt{\frac{2}{n}} \quad (73)$$

$$c^\top x + 8 = 4 \quad (74)$$

$$c^\top x = -4 \quad (75)$$

$$n(ca) = -4 \quad (76)$$

$$c = \frac{-4}{an} \quad (77)$$

$$c = \mp \frac{4}{\sqrt{2n}} \quad (78)$$

Please noted that here c and a have inverse signs.

c) KKT Condition For original problem:

$$\min f(x) \quad (79)$$

$$h(x^*) = 0 \quad (80)$$

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0 \quad (81)$$

KKT Condition For In-Eq problem:

$$g(x^*) = \frac{1}{2} \|h(x^*)\|^2 \leq 0 \quad (82)$$

$$\mu^* \geq 0 \quad (83)$$

$$\nabla f(x^*) + \mu^* h(x^*) = 0 \quad (84)$$

$$\frac{1}{2} \mu^* \|h(x^*)\|^2 = 0 \quad (85)$$

If $\mu^* = 0$:

$$\nabla f(x^*) = 0 \quad (86)$$

$$h(x^*) \text{ may not be } 0 \quad (87)$$

So $\mu^* > 0$:

$$\|h(x^*)\|^2 = 0 \quad (88)$$

$$\nabla g(x^*) = \|h(x^*)\| = 0 \quad (89)$$

$$h(x^*) = 0 \quad (90)$$

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.