ORFE 522: Linear and Nonlinear Optimization
Homework 3

Zachary Hervieux-Moore

Friday 18$^{\text{th}}$ November, 2016

Exercise 1: Consider the problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \text{ s.t.} g_j(x) \leq 0, \; j = 1, \dots, m$$

where f, g_j are continuously differentiable. Suppose that x^* is a local optimal solution and $\nabla g_1(x^*), \dots, \nabla g_m(x^*)$ are linearly independent. Prove that x^* is a global optimal solution to the linearized problem:

$$\begin{gathered} \min_{\mathbf{x}} f(\mathbf{x}^*) + \nabla f(x^*)^T(x - x^*) \\ \text{s.t. } g_j(x^*) + \nabla g_J(x^*)^T(x - x^*) \leq 0, \; j = 1, \dots, m \end{gathered}$$

Answer: We will split it into two cases. First, when the optimal solution x^* has no active contraints. Then when x^* has active constraints.

If there are no active constraints, then $\nabla f(x^*) = 0$ and the linearized problem reduces to

$$\begin{gathered} \min_{\mathbf{x}} f(\mathbf{x}^*) \\ \text{s.t. } g_j(x^*) + \nabla g_j(x^*)^T(x - x^*) \leq 0, \; j = 1, \dots, m \end{gathered}$$

Picking $x = x^*$ clearly satisfies the constraints since $g_j(x^*) \leq 0$ as x^* is a solution to the first problem. It is also a global optimum since it is constant.

Now consider the case when there are active constraints. That is, $g_i(x^*) = 0$ for some i's. Since we have linear independence of the $\nabla g_j(x^*)$, we know the KKT conditions are satisfied. Thus,

$$-\nabla f(x^*) = \sum_{j=1}^{m} \mu_j \nabla g_j(x^*)$$

With $\mu_j \geq 0$. Then the new linearized problem can equivalently be written as

$$f(\mathbf{x}^*) - \sum_{j=1}^{m} \mu_j \nabla g_j(x^*)^T(x - x^*)$$

Since, by complementary slackness, $\mu_j = 0$ for non-active constraints, we can ignore them. For active constraints $\mu_i \geq 0$. Thus, since we are subtracting the sum, we only need to worry about

$$\nabla g_j(x^*)^T(x - x^*) > 0$$

Since this would make the objective smaller.

Note that if $\nabla g_i(x^*)^T(x - x^*) \leq 0$ then picking $x = x^*$ would minimize this addition (since we are subtracting negative values). Now suppose that $\nabla g_i(x^*)^T(x - x^*) > 0$. Then, since these are active constraints, $g_i(x^*) = 0$ and the inequalities becomes $\nabla g_i(x^*)^T(x - x^*) \leq 0$ which does not meet our assumption. Thus, we cannot have $\nabla g_i(x^*)^T(x - x^*) > 0$ for active constraints and $\mu_j = 0$ for non-active constraints. Thus, $x = x^*$ becomes a global optimal to remove any of the positive terms that result from the sum. Also, when $x^* = x$, the contraints are satisfied due to x^* satisfying the original problem.

Exercise 2: Consider the function

$$f(x, y, z) = 2x^2 + xy + y^2 + yz + z^2 - 6x - 7y - 8z - 9$$

1) By using the first-order necessary conditions, find the candidate minimum points of $f(x, y, z)$.

2) Verify using the second-order sufficient conditions whether those points are local minimum.

3) Find a global minimum point.

4) Use CVX (or any other optimization solver) to verify your results (please print your codes on a paper, same for the next question). Here you may use the following form when inputting into CVX

$$f(x, y, z) = x^2 + \frac{1}{2}y^2 + \left(x + \frac{1}{2}y\right)^2 + \left(z + \frac{1}{2}y\right)^2 - 6x - 7y - 8z - 9$$

Answer:

1) We calculate the gradient and set it to zero to find the candidate points,

$$\nabla f(x, y, z) = \begin{bmatrix} 4x + y - 6 \\ x + 2y + z - 7 \\ y + 2z - 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently,

$$\begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix}$$

Solving this linear equation yields the candidate point $(\frac{6}{5}, \frac{6}{5}, \frac{17}{5})$.

2) Calculating the Hessian,

$$H(f) = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Applying Sylvester's criterion and calculating the leading principal minors. $M_1 = 4 > 0$, $M_2 = 7 > 0$, and $M_3 = 10 > 0$. Therefore the Hessian is positive definite for all (x, y, z) and hence the candidate point is a local minimum.

3) Since there is only one local minimum, it must be a global minimum. Thus the point $(\frac{6}{5}, \frac{6}{5}, \frac{17}{5})$ is a global minimum.

4) The MATLAB CVX code below generates a solution of (1.2000, 1.1999, 3.4001) which is quite close to the analytical answer.

```
clear;
clc;

cvx_begin
    variable x(3);
    minimize( x(1)^2 +1/2*x(2)^2 + (x(1)+1/2*x(2))^2 +(x(3)+1/2*
        x(2))^2 - 6*x(1) - 7*x(2) - 8*x(3) - 9 );
cvx_end

x
```

Exercise 3: We want to build a cylinder with the maximum volume, with its surface area no larger than C. The decision variables are: r (the radius of the base) and h (height). Then the optimization problem is:

$$\begin{aligned} &\max_{r,h} \pi r^2 h \\ \text{s.t. } &2\pi r^2 + 2\pi r h \leq C \\ &r, h \geq 0 \end{aligned}$$

1) Does there exist an optimal solution? Why?
2) Is this a convex problem?
3) Find the set of optimal solutions using KKT conditions.

Answer:

1) Let $z = rh$. Then the problem becomes

$$\begin{aligned} &\max_{r,z} \pi r z \\ \text{s.t. } &2\pi r^2 + 2\pi z \leq C \\ &r, h \geq 0 \end{aligned}$$

   Thus, the feasibility set it $(r, z) \in [0, \sqrt{C/2\pi}] \times [0, C/2\pi]$ which is clearly compact. Thus, since $\pi r z$ is continuous, a maximum is achieved on this compact set. Note that $h = r/z$ is well defined because $z \neq 0$. Otherwise, if $z = 0$, then the volume is 0 which is clearly not a maximum.

2) This is not a convex problem as the Hessian of the objective is not positive semidefinite

$$H(f) = \begin{bmatrix} 2\pi h & 2\pi r \\ 2\pi r & 0 \end{bmatrix}$$

   Using Sylvester's criterion, the leading principal minors are $M_1 = 2\pi h \geq 0$ and $M_2 = -4\pi r^2 \leq 0$. Thus this is not positive semidefinite and so not convex.

3) The KKT conditions give

$$\nabla f = \begin{bmatrix} 2\pi rh \\ \pi r^2 \end{bmatrix} = \mu_1 \begin{bmatrix} 4\pi r + 2\pi h \\ 2\pi r \end{bmatrix} + \mu_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mu_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

We also have the complementary slackness conditions

$$\begin{aligned} \mu_1(2\pi r^2 + 2\pi rh - C) &= 0 \\ -\mu_2 r &= 0 \\ -\mu_3 h &= 0 \end{aligned}$$

Thus, noting the fact that if r or $h = 0$, then the volume is 0 (a minimum), we must have that $\mu_2 = \mu_3 = 0$. From stationarity equation, if $\mu_1 = 0$, then this implies that $r = 0$. So we must have $\mu_1 > 0$ and $2\pi r^2 + 2\pi rh - C = 0$. Rearranging some equations, we get

$$\begin{aligned} \frac{r}{2} &= \mu_1 \\ \implies 2\pi rh &= \frac{r}{2}(4\pi r + 2\pi h) \iff h = 2r \end{aligned}$$

Putting this into the constraint equation

$$6\pi r^2 = C \iff r = \sqrt{\frac{C}{6\pi}}$$

Thus, the set of optimal solutions is $\mu_1 = \frac{1}{2}\sqrt{\frac{C}{6\pi}}$, $\mu_2 = \mu_3 = 0$, $r = \sqrt{\frac{C}{6\pi}}$, and $h = 2\sqrt{\frac{C}{6\pi}}$

Exercise 4: Either prove or find a counterexample for each of the following statements (you can assume all the functions are second order continuously differentiable):

1) If $f(x)$ is convex, $g(x)$ is convex, then $f(g(x))$ is convex.
2) If $f(x)$ is convex and nondecreasing, $g(x)$ is convex, then $f(g(x))$ is convex.
3) If $f(x)$ is concave and nonincreasing, $g(x)$ is convex, then $f(g(x))$ is convex.
4) If $f(x)$ is increasing and non-negative, then $xf(x)$ is convex on $x \geq 0$.

Answer:

1) False. Consider $f(x) = e^{-x}$ and $g(x) = x^2$ which are both convex. Then $f(g(x)) = e^{-x^2}$ and the second derivative is $e^{-x^2}(4x^2 - 2)$ which is negative at 0 and hence not convex.

2) True. Taking the derivatives via chain rule, we get that the second derivative of $f(g(x))$ is

$$f''(g(x))g'(x)^2 + f'(g(x))g''(x)$$

Since $f(x)$ is convex and nondecreasing, then $f''(x) \geq 0$ and $f'(x) \geq 0$. Since $g(x)$ is convex, then $g''(x) \geq 0$ and $g'(x)^2 \geq 0$ since it is squared. Thus, all the terms are nonnegative and so $f''(g(x))g'(x)^2 + f'(g(x))g''(x) \geq 0$ and hence $f(g(x))$ is convex.

3) False. Using the same decomposition as last part. We get that

$$f''(g(x))g'(x)^2 + f'(g(x))g''(x) \leq 0$$

Since $f''(x) \leq 0$ as it is concave and $f'(x) \leq 0$ as it is nonincreasing. Also $g(x)$ is convex, so $g''(x) \geq 0$ and $g'(x)^2 \geq 0$ since it is squared. Putting it all together, we get that that $f(g(x))$ is concave. A counterexample is $f(x) = -e^x$ and $g(x) = x^2$. Clearly, $f(x)$ is concave and nonincreasing. Then $f(g(x))'' = -2e^{x^2}(2x^2+1)$ which is negative every where and so concave.

4) False. Consider $f(x) = 1 - \frac{1}{(x+1)^2}$ on $\mathbb{R}^+$. Then this is non-negative as it starts at 0 and tends to 1. It is also increasing since the first derivative is $f'(x) = \frac{2}{(x+1)^3}$ which is positive for $x \geq 0$. However, the second derivative of $xf(x)$ is $-\frac{2(x-2)}{(x+1)^4}$ which is negative on $x > 2$ and so not convex. If you wish to consider the entire real line, define $f(x) = 0$ for $x \in R^-$.