Math 408 Homework Set 6 Solutions

(1) Show that the functions

$$f(x_1,x_2)=x_1^2+x_2^3, \quad \text{and} \quad g(x_1,x_2)=x_1^2+x_2^4$$

both have a critical point at $(x_1,x_2)=(0,0)$ and that their associated Hessians are positive semi-definite. Then show that $(0,0)$ is a local (global) minimizer for g and not for f.

Solution
Both f and g are completely separable, i.e. they are the sum of functions of the components and have the form $h(x)=\sum_{i=1}^n h_i(x_i)$. The origin is the unique critical point for both functions. However,

$$\nabla^2 f(x_1,x_2)=\begin{bmatrix} 2 & 0 \\ 0 & 6x_2 \end{bmatrix} \qquad \nabla^2 g(x_1,x_2)=\begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}.$$

Clearly, $\nabla^2 f$ is not positive semi-definite for $x_2<0$, so f is not minimized at the origin, while $\nabla^2 g$ is everywhere positive definite away from the origin and positive semidefinite at the origin. Thus, f has no local (global) optima, while the origin is a global minimizer of g.

(2) Find the local minimizers and maximizers for the following functions if they exist:
 (a) $f(x)=x^2+\cos x$
 (b) $f(x_1,x_2)=x_1^2-4x_1+2x_2^2+7$
 (c) $f(x_1,x_2)=e^{-(x_1^2+x_2^2)}$
 (d) $f(x_1,x_2,x_3)=(2x_1-x_2)^2+(x_2-x_3)^2+(x_3-1)^2$

Solution
 (a) $x=0$ is local (global) minimizer;
 (b) $(x_1,x_2)^T=(2,0)^T$ is local (global) minizer;
 (c) $(x_1,x_2)=(0,0)$ is local (global) maximizer;
 (d) $(x_1,x_2,x_3)=(\frac{1}{2},1,1)$ is local (global) minimizer.

(3) Compute the directional derivative for each of the following functions at the origin.
 (a) $f(x)=\max\{0,x\}$
 (b) $f(x)=\max\{-x,\ 2x\}$
 (c) $f(x_1,x_2)=|x_1|-|x_2|$

Solution:
 (a)

$$f'(x;d)=\begin{cases} d & , x>0 \text{ or } (x=0 \text{ and } d>0), \\ 0 & , x<0 \text{ or } (x=0 \text{ and } d\le 0). \end{cases}$$

 (b)

$$f'(x;d)=\begin{cases} 2d & , x>0 \text{ or } (x=0 \text{ and } d>0), \\ -d & , x<0 \text{ or } (x=0 \text{ and } d\le 0). \end{cases}$$

(c)

$$f'(x;d) = \begin{cases} d_1 - d_2 & , (x_1 > 0, x_2 > 0) \wedge (x_1 = 0, x_2 > 0, d_1 \geq 0) \wedge (x_1 > 0, x_2 = 0, d_2 \geq 0) \wedge (x_i = 0, d_i \geq 0, i = 1, 2), \\ d_1 + d_2 & , (x_1 > 0, x_2 < 0) \wedge (x_1 = 0, x_2 < 0, d_1 \geq 0) \wedge (x_1 > 0, x_2 = 0, d_2 \leq 0) \wedge (x_i = 0, d_1 \geq 0, d_2 \leq 0), \\ -d_1 - d_2 & , (x_1 < 0, x_2 > 0) \wedge (x_1 = 0, x_2 > 0, d_1 \leq 0) \wedge (x_1 < 0, x_2 = 0, d_2 \geq 0) \wedge (x_i = 0, d_1 \leq 0, d_2 \geq 0), \\ -d_1 + d_2 & , (x_1 < 0, x_2 < 0) \wedge (x_1 = 0, x_2 < 0, d_1 \leq 0) \wedge (x_1 > 0, x_2 = 0, d_2 \leq 0) \wedge (x_i = 0, d_i \leq 0, i = 1, 2). \end{cases}$$

(4) Show that the function $f(x) := \frac{1}{2}(\max\{0,\ x\})^2$ is differentiable at the origin and give its derivative.

Solution: Simply note that f is differentiable at any point $x \neq 0$ so $f'(x) = 0$ for $x > 0$ and $f'(x) = x$ for $x > 0$. The only difficulty occurs at $x = 0$. In this case just observe that

$$0 = \lim_{\Delta x \downarrow 0} \frac{1}{2} \Delta x = \lim_{\Delta x \downarrow 0} \frac{f(\Delta x) - f(0)}{\Delta x}$$

and

$$0 = \lim_{\Delta x \uparrow 0} \frac{f(\Delta x) - f(0)}{\Delta x},$$

so that the derivative at $x = 0$ is zero. Hence, $f'(x) = \max\{0, x\}$.

(5) Let $C \subset \mathbb{R}^n$ and $x \in C$ and recall the definition of the tangent cone to C at x:

$$T_C(x) := \left\{ u \;\middle|\; \exists\ \{x^\nu\} \subset C,\ x^\nu \to x,\ t_\nu \downarrow 0, \quad \text{with}\ \ t_\nu^{-1}(x^\nu - x) \to u \right\}.$$

(a) Let $\mathbb{B}_2 = \{u \,|\|u\|_2 \leq 1\}$. Show that for all $u \in \mathbb{B}_2$ with $\|u\|_2 = 1$,

$$T_{\mathbb{B}_2}(u) = \left\{ v \;\middle|\; u^T v \leq 0 \right\}.$$

Solution: Set $c(x) := \|x\|_2^2 = \sum_{i=1}^n x_i^2$ so that $\nabla c(x) = x$. In this case the set C is $\mathbb{B}_2 = \{x \,|c(x) \leq 1\}$. Since this representation for $\mathbb{B}_2$ satisfies the LICQ for all $x \in \mathbb{B}_2 \setminus \{0\}$, we know that for every $u \in \mathbb{B}_2$ with $\|u\|_2 = 1$ satisfies $T_{\mathbb{B}_2}(u) = \left\{ v \;\middle|\; \langle u, v \rangle = \nabla c(u)^T v \leq 0 \right\}$.

(b) Consider the continuous function

$$f(x) := \begin{cases} -\sqrt{\|x\|_2^2 - 1} & , \text{if } \|x\|_2 \geq 1, \text{ and} \\ 0 & , \text{if } \|x\|_2 < 1. \end{cases}$$

Obviously, $\mathbb{B}_2 = \operatorname{argmin} \{ f(x) \,|x \in \mathbb{B}_2 \}$, since f is identically zero on $\mathbb{B}_2$. Let $\|u\|_2 = 1 = \|v\|_2$ with $u^T v = 0$ so that $v \in T_{\mathbb{B}_2}(u)$. Show that $f'(u; v)$ exists with $f'(u; v) = -1$, where

$$f'(u; v) := \lim_{t \downarrow 0} \frac{f(u + tv) - f(u)}{t}\ .$$

Solution: $f'(u;v) := \lim_{t \downarrow 0} \frac{-\sqrt{\|u+tv\|_2^2 - 1} + \sqrt{\|u\|_2^2 - 1}}{t} = \lim_{t \downarrow 0} \frac{-\sqrt{t\langle u,v\rangle + t^2\|v\|_2^2}}{t} = \lim_{t \downarrow 0} -\frac{|t|}{t} = -1.$

(c) Let $h : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and let $S \subset \mathbb{R}^n$. Recall from Theorem 4.2 in Chapter 4 of the course notes that if $\bar{x} \in \operatorname{argmin}\{h(x) \mid x \in S\}$, then $h'(\bar{x}; d) \geq 0$ for all $d \in T_S(\bar{x})$. Does this result contradict your finding in part (5b)? If not, why not?

Solution: On the surface it would appear that the optimization problem above violates the theorem since the theorem states that $f'(x : d) \geq 0 \,\forall\, d \in T_\Omega(x)$ whenever $x \in \operatorname{argmin}\{f(x) \mid x \in \Omega\}$. However, the theorem requires that f be continuously differentiable at the point x. The function defined above is not differentiable on the boundary of the unit ball $\mathbb{B}_2$.

(6) Show that the representation of the set $\Omega := \{x \in \mathbb{R}^2 | -x_1^3 \leq x_2 \leq x_1^3\}$ is not regular at the origin. Can you suggest an alternative representation that is regular at the origin?

Solution: $T_\Omega(0,0) = \{0\} \times \mathbb{R}_+$ while

$$\left\{d \,\middle|\, \nabla f_1(0,0)^T d \leq 0 \text{ and } \nabla f_1(0,0)^T d \leq 0\right\} = \{0\} \times \mathbb{R}\ ,$$

where $f_1(x_1, x_2) = -(x_1^3 + x_2)$ and $f_2(x_1, x_2) = x_2 - x_1^3$.

(7) Let Ω be given the representation $\Omega := \{x \in \mathbb{R}^2 \,|x_2 \leq 0,\ -x_2 \leq 0\}$ and consider the optimization problem $\min\{x_1^2 \,|x \in \Omega\}$. Show that the unique global minimzer of this problem satisfies the MFCQ but not the LICQ. Also, compute the set of KKT multipliers for this global solution.

Solution: Observe that $\Omega = \left\{(x_1, x_2)^T \,|x_2 = 0\right\}$ so that $(x_1, x_2)^T = (0,0)^T$ is the global optimal solution. Let $c_1(x) = x_2$ and $c_2(x) = -x_2$ so that $\nabla c_1(x) = (0,1)^T$ and $\nabla c_2(x) = (0,-1)^T$. Since $\nabla c_1(x)$ and $\nabla c_2(x)$ are nowhere linearly independent, the LICQ cannot hold anywhere. On the other hand, the MFCQ is satisfied everywhere.

(8) Locate all of the KKT points for the following problems. Can you show that these points are local solutions? Global solutions?

(a)

$$\begin{array}{ll} \text{minimize} & e^{(x_1 - x_2)} \\ \text{subject to} & e^{x_1} + e^{x_2} \leq 20 \\ & 0 \leq x_1 \end{array}$$

Solution: This is convex problem so any local solution is a global solution. Obviously, we wish to make x_1 as small as possible and x_2 as big as possible. Hence, we must have $x_1 = 0$ which gives the solution $(x_1, x_2) = (0, \ln(20-1))$. By plugging this solution into the KKT conditions, we obtain the multipliers $(y_1, y_2) = (1,2)/19$.

(b)

$$\begin{array}{ll} \text{minimize} & e^{(-x_1 + x_2)} \\ \text{subject to} & e^{x_1} + e^{x_2} \leq 20 \\ & 0 \leq x_1 \end{array}$$

Solution: Here we want to make x_1 as big as possible and x_2 as small as possible. By fixing $x+1$ at zero and sending x_2 to $-\infty$, the constraints are satisfied and the objective goes to zero. Hence, no solution exists and the optimal value is 0.

(c)

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 - 4x_1 - 4x_2 \\ \text{subject to} & x_1^2 \leq x_2 \\ & x_1 + x_2 \leq 2 \end{array}$$

Solution: This is a convex optimization problem and so any KKT point will give global optimality. Check that $(x_1, x_2) = (1,1)$ and $(y_1, y_2) = (0,2)$ is a KKT pair for this problem.

(d)

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|x\|_2^2 \\ \text{subject to} & Ax = b \end{array}$$

where $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$ satisfies $\mathrm{Nul}\,(A^T) = \{0\}$.

Solution: This is a convex problem so $\bar{x}$ is a solution if and only if there is a $\bar{y}$ such that $(\bar{x}, \bar{y})$ is a KKT pair for this problem. The Lagrangian is $L(x, y) = \frac{1}{2}\|x\|_2^2 + y^T(b - Ax)$. The KKT conditions are $A\bar{x} = b$ and $\bar{x} = A^T\bar{y}$. Hence $b = A\bar{x} = AA^T\bar{y}$. Since $\mathrm{Nul}\,(A^T) = \{0\}$, $\mathrm{Nul}\,(AA^T) = \{0\}$ so that the matrix AA^T is invertible. Consequently, $\bar{y} = (AA^T)^{-1}b$ and $\bar{x} = A^T\bar{y} = A^T(AA^T)^{-1}b$.

(9) Suppose $\Omega = \{x\,;\, Ax \le b,\ Ex = h\}$ where $A \in \mathbb{R}^{m\times}$, $E \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^m$, and $h \in \mathbb{R}^k$.

(a) Given $x \in \Omega$, show that

$$T_\Omega(x) = \{d \,:\, A_{i\cdot}d \le 0 \text{ for } i \in I(x),\ Ed = 0\},$$

where $A_{i\cdot}$ denotes the ith row of the matrix A and $I(x) = \{i\ \ A_{i\cdot}x = b_i\}$.

Solution: It was shown in class (see page 3 of the course notes *Optimality Conditions for Constrained Problems*) that

$$(\spadesuit) \qquad T_\Omega(x) \subset \{d \,:\, A_{i\cdot}d \le 0 \text{ for } i \in I(x),\ Ed = 0\},$$

so we need only show the reverse inclusion. Let $x \in \Omega$ and d be and element of the set of the right hand side of ($\spadesuit$). Note that by continuity there is a $\bar{t} > 0$ such that $x + td \in \Omega$ for all $0 \le t \le \bar{t}$. Hence $d \in T_\Omega(x)$.

(b) Given $x \in \Omega$, show that every $d \in T_\Omega(x)$ is a feasible direction for Ω at x.

Solution: This is what we showed in the answer to the previous question.

(c) Note that parts (a) and (b) above show that

$$T_\Omega(x) = \bigcup_{\lambda > 0} \lambda(\Omega - x)$$

whenever Ω is a convex polyhedral set. Why?

Solution: Because every polyhedral convex set canb be given a representation

$$\Omega = \{x\,;\, Ax \le b,\ Ex = h\}\ .$$